THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR .... Arthur Donald Stewart Steckhahn ....

TITLE OF THESIS .... Multivariable System Stability Via Volterra ....

.... Series ....

....

DEGREE FOR WHICH THESIS WAS PRESENTED .... Master of Science ....

YEAR THIS DEGREE GRANTED .... 1975 ....

THE UNIVERSITY OF ALBERTA

MULTIVARIABLE SYSTEM STABILITY VIA VOLTERRA SERIES

by

ARTHUR DONALD STEWART STECKHAHN

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE.

DEPARTMENT: ELECTRICAL ENGINEERING

EDMONTON, ALBERTA

FALL 1975

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled ..Multivariable System Stability..................

...................Via Volterra Series.................................

....

submitted by ....

in partial fulfilment of the requirements for the degree of

Master of Science

in Electrical Engineering.

ABSTRACT

The stability analysis of a class of non-linear systems is presented in this thesis. The class of systems are those which can be represented by vector matrix state equations, are analytic and have a non-linearity (which passes through the origin), in the feed-back loop. Two types of systems are considered, the first is a continuous system, the second has the same system equation but is considered from the sampled data point of view; with and without zero order hold. Non-zero initial conditions are considered from the outset and inter-sample stability in the case of the sampled data systems is considered. The determination of the input-output stability bounds is approached via the Volterra series, whose region of convergence and uniqueness is assured by satisfying the Banach contraction-mapping principle. Determination of the contraction constant is achieved via the Frechet derivative.

ACKNOWLEDGEMENTS

The author wishes to thank his supervisor Dr. G. S. Christensen for his guidance and help throughout the course of this research and his help structuring the draft of the thesis.

The financial assistance of the University of Alberta (1970-71) is gratefully acknowledged.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Modern technology has given rise to today's complex systems many of which are only describable in terms of non-linear vector differential equations; the suggestion that these systems can be analysed for stability by the application of straight-forward methods, applicable to simpler systems, is not always true. Dimensionality can produce problems which had not been considered at the outset; thus more sophisticated techniques must be used in the determination of system stability.

What is system stability, and how is this stability determined?

These questions are of paramount importance to the systems engineer seeking the region of stability meaningful to his particular problem.

The concept of stability of a linear system with constant coefficients is basic to control engineering, stability as defined by Bower and Schultheiss, [1]† is that a linear system is stable if and only if its output response to every bounded input remains bounded. Thus a linear system can be theoretically stable for any input regardless of size. Such is not the case for non-linear systems.

The stability of non-linear systems is often a local concept and is possibly some function of the input. The definitions of non-linear system stability existing in the literature are many, some of which are referenced below. The list is by no means exhaustive.

(1) Kalman & Bertram [2] have given eight definitions of stability.

(2) Antosiewicz [3] has nine definitions.

(3) Ingwerson [4] gives twenty definitions.

---

† Numbers placed within square brackets denote references.

The definition of stability to be adopted here is that of Zadeh [5], that a bounded set of inputs gives rise to a bounded set of outputs, hereafter called B.I.B.O. stability. In non-linear systems, the determination of the exact stability boundary is extremely difficult and approximate techniques have, since the early 1950s, become more apparent in the literature. It is instructive to trace the development of what is now called "stability analysis via the Volterra series".

An idea was conceived by the mathematician Volterra [6] of creating a general theory of functions which depend upon a continuous set of values of another function. This theory is now called the "Theory of Functionals", a name coined by Hadamard. The theory lay virtually dormant until Wiener (1942) utilized it in his research, Barrett utilizing the functional power series (due to Volterra) and the work accomplished by Wiener, discussed a method of expanding the input - output relationship of a non-linear system in terms of a functional power series.

It was not until the late 1950s and early 1960s however, that the potential of the Volterra series was realised.

Brilliant [7] utilizing Wiener's effort specified the conditions for an analytic system. Flake [8] showed that the Volterra series could be used to relate the input to the output of a non-linear system. Barrett [9] demonstrated Flake's concept by using Duffing's Equation and by utilizing the analytic concept of Brilliant he proved that the resulting Volterra series was convergent. The convergence properties were directly related to boundedness and stability of the system.

Christensen [10] then proved that the convergence properties of the Volterra series were obtainable via the Banach contraction principle. Then, by further application of the contraction principle he proved that a stable region could be found in the Banach space under consideration, this space contained the solutions of the non-linear system which itself exhibited B.I.B.O. stability in the region so determined. Corroborative evidence was now established between the methods displayed by Barrett and Christensen.

Addition of initial conditions, uniqueness of the Volterra series and a larger region of convergence were contributed by Trott and Christensen [34], [33] and [11] respectively.

Applications to discrete data systems was made by Rashed [12] and Rao [13].

To date the most significant results have been achieved with systems that are reducable to the form shown below.

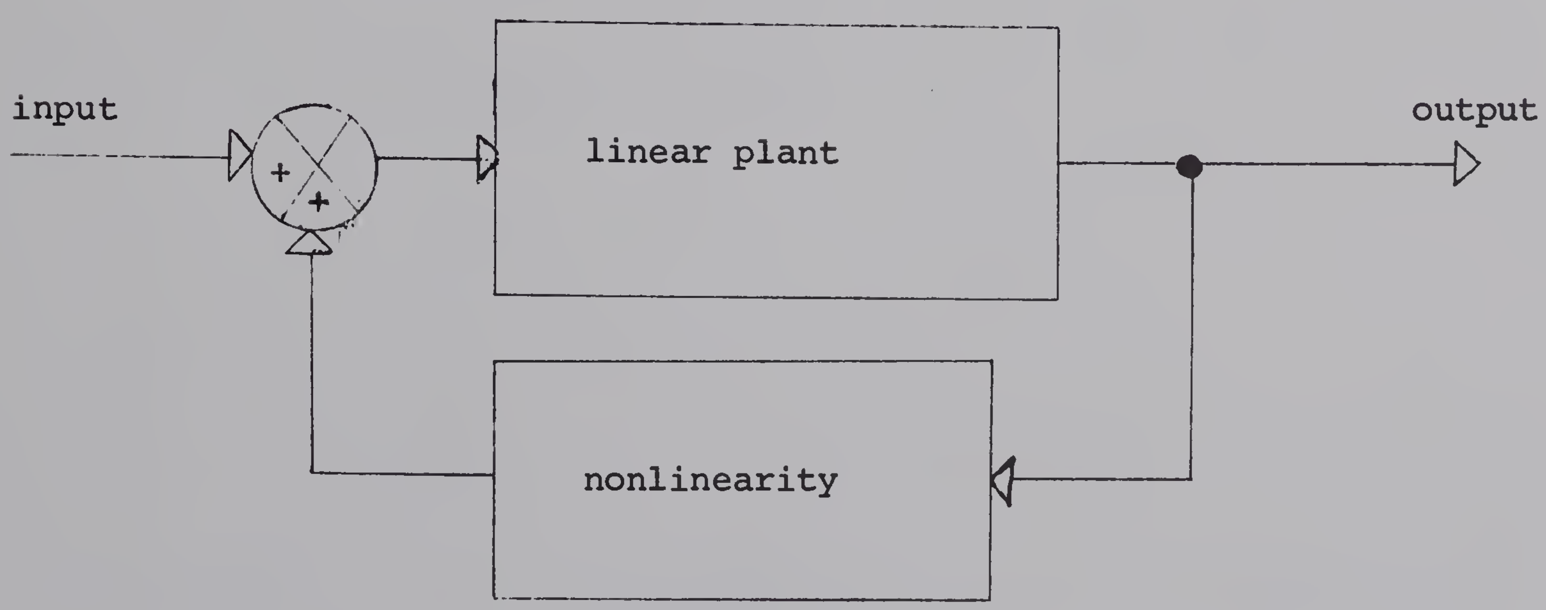

Figure 1: Single - input/single - output systems.

Thus the extension of the work of Christensen, Trott, Rao, Rashed to the multivariable system is the next logical step.

This thesis will extend their results to systems expressible as non-linear vector differential equations. The extension will be made via the contraction mapping principle utilizing the Frechet differential and state-space matrix methods, [15], [24], [26]. The class of systems considered here will contain a time invariant non-linearity passing through the origin.

Chapter II will introduce the notation and mathematics required in the sequel, theorems current to this thesis will be stated and proofs shown where necessary.

In Chapter III non-linear system equations will be examined, the region of stability determined by (1) reduction of the system equation to standard matrix form, and (2) by reduction of the system equation by Heaviside's decomposition.

The methods of Chapter III will be extended in Chapter IV to cover discrete data systems.

CHAPTER II

MATHEMATICAL DERIVATIONS

2.0 Introduction.

The mathematics necessary for the derivation of the material required in the sequel is presented in this chapter. The object is not to present a mathematical treatise, but rather the material necessary for the maintenance of continuity once the development of the core of the thesis is underway. Functions representative of practical systems are considered and the space on which they operate is a Banach space.

Matrices and norms have an important role in the ensuing mathematical development. It is impossible to study multi-dimensional systems in a meaningful fashion without matrices, while the norm is one of the foundation stones of functional analysis.

The Frechet (F) derivative is approached via the F differential which is approached via the Gateaux (G) differential. A pictorial representation of the route taken in the development of the F derivative is shown in Fig. 2, Ortega and Rheinbolt, [19].

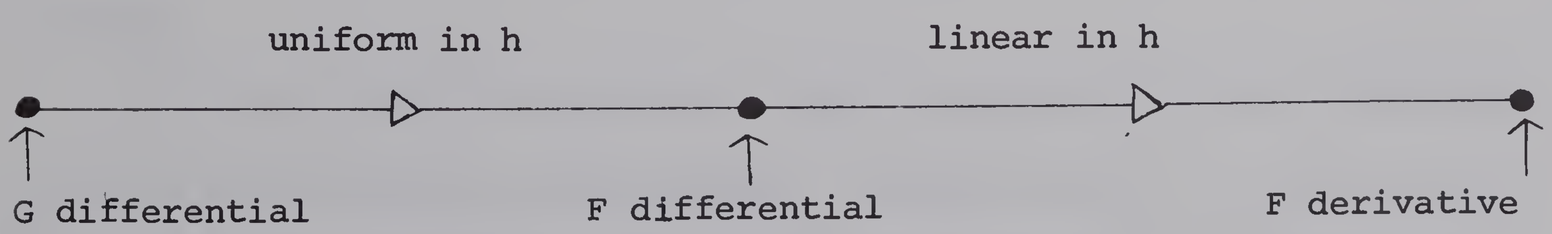

Figure 2: Relationship between "F" differential and the "F" derivative

A transformation which takes a domain of the Banach space $\underline{X}$ into itself is considered and a contraction of the transformation will exist in the domain, if the following conditions are satisfied Nashed, [18]:

(1) The domain is invariant under the transformation.

(2) The norm of the F derivative is less than one.

Some of the questions to be answered in this chapter are:

(1) What are the existence conditions for the F derivative.

(2) What conditions must exist for an operator (mapping function) to be continuous.

(3) Why use the F derivative.

2.1 Functional Representation.

A control system can be viewed as a transformation which maps points of the input space into the output space. The process is iterative. The existence of a solution and the method of obtaining it will, for the system in question, depend upon the transformation.

Many existence theorems, in functional analysis, can be proved with the aid of the Banach contraction - mapping principle. Clearly the determination of the existence of a solution for the system in question via the contraction - mapping principle is relevant in this context.

Since the Banach contraction - mapping principle forms an integral part of the mathematical development, a space conducive to the use of the principle is required. This space, as stated in the introduction to this chapter, is a Banach space.

<u>Definition 1</u>: A "transformation", (denoted by the symbol $\underline{\Gamma}$ - Gamma), is a function defined on a vector space $\underline{X}$ which takes its values in another vector space $\underline{Y}$. An element of $\underline{Y}$ is called an image of $\underline{x} \in \underline{X}$ under $\underline{\Gamma}$.

<u>Definition 2</u>: An "operator" is a transformation whose domain and range lie in the same space.

An operator will also be denoted by $\underline{\Gamma}$.

<u>Definition 3</u>: A "space" R is a set of elements f, g, h, ----, these elements may be real or complex numbers, vectors, matrices or functions of one or more variables.

The spaces considered here are linear.

<u>Definition 4</u>: A "linear space" is one in which the elements of the space obey the laws of linearity.

<u>Definition 5</u>: A "vector space", $\underline{X}$, consists of a set of quantities $\underline{x}$, $\underline{y}$, ---, (called vectors), and two operations.

(1) + (plus) defined as pointwise addition.

(2) • (multiplication) defined as scalar multiplication.

The space is closed under + and • if the following axioms and their combinations are satisfied:

(a) $(\underline{x} + \underline{y}) = \underline{z}, \quad \forall\ \underline{x}, \underline{y}, \underline{z}, \in \underline{X}$

(b) $c\underline{x} \in \underline{X}, \quad \forall\ \underline{x} \in \underline{X}$ and all c (2.1)

(c) $c\underline{\theta} = \underline{\theta}$ is the null vector.

A linear space that has a finite algebraic base is said to be finite dimensional.

<u>Definition 6</u>: The "dimension n" of a vector space is the number of elements in a basis. Every non-zero finite dimensional vector space has a basis.

If $\underline{X}$ is a finite dimensional vector space then every element of $\underline{X}$ can be represented in the following form:

$$\underline{x} = \sum_{i=1}^{n} \gamma_i x_i \tag{2.2}$$

where the x_is form the algebraic basis.

<u>Defintion 7</u>: The "norm" ($\|\cdot\|$) of a vector is a mapping from the vector space $\underline{X}$ to $R=E^1$ (the space of real numbers), and is said to be a functional of $\underline{X}$. To be a norm, the functional must satisfy the following axioms [21]-[26]:

(a) $\|\underline{x}\| \geq 0 \qquad \forall\ \underline{x} \in \underline{X}$

$\|\underline{x}\| = 0 \qquad \text{iff } \underline{x} = \underline{\theta}$

(b) $\|\underline{x} + \underline{y}\| \leq \|\underline{x}\| + \|\underline{y}\| \qquad \forall\ \underline{x}, \underline{y} \in \underline{X}$ (2.3)

(c) $\|c\underline{x}\| \leq |c|\ \|\underline{x}\| \qquad \forall\ \underline{x} \in \underline{X} \text{ and } \forall\ c \in E^1$

(d) $\Big|\|\underline{x}\| - \|\underline{y}\|\Big| \leq \|\underline{x} - \underline{y}\| \qquad \forall\ \underline{x}, \underline{y} \in \underline{X}$

By inspection of (2.3) it can be seen that a norm is not a linear functional, but because of (2.3(d)) it is a continuous functional.

Thus the norm is a continuous functional, but it is not necessarily linear. Loosely speaking, the norm of a vector is a measure of its "length"; similarly the norm of a matrix $\underline{A}$ is a measure of the "size of $\underline{A}$".

The norm of a matrix vector product occurs frequently throughout the mathematical development, thus (2.3) will be augmented by an additional condition which connects matrix and vector norms.

To a given vector norm, a compatible matrix norm is defined by:

$$\|\underline{A}\| = \sup_{\|\underline{x}\| \neq 0} \frac{\|\underline{A}\,\underline{x}\|}{\|\underline{x}\|} \tag{2.4}$$

thus $$\|\underline{A}\,\underline{x}\| \leq \|\underline{A}\|\;\|x\|.$$

<u>Proof</u>:

$$\underline{y} = \underline{A}\,\underline{x} \quad \forall\ \underline{x}, \underline{y} \in \underline{X}$$

where $\underline{A} = (a_{ij}) \qquad i,j = 1, 2, ----, n$

$\underline{X} = (x_1, x_2, -----, x_n)$

$\underline{y} = (y_1, y_2, ------, y_n)$

thus $$y_i(t) = \sum_{j=1}^{n} a_{ij}x_j(t) \qquad i = 1, 2, ----, n,$$

therefore,

$$\left|y_i(t)\right| \leq \sum_{j=1}^{n} |a_{ij}|\,|x_j(t)|. \tag{2.5}$$

The supremum norm - details of which are contained in [23] - is used throughout.

The definitions of matrix and vector norms respectively are:

$$\|\underline{A}\| = \sup_i \sum_{j=1}^{n} |a_{ij}| \tag{2.6}$$

$$\|\underline{x}\| = \sup_i \sup_t \left| x_i(t) \right|. \tag{2.7}$$

Thus (2.5) can be rewritten as

$$\sup \left| y_i(t) \right| \leq \sup_i \left(\sum_{j=1}^{n} |a_{ij}| \sup_t \left| x_j(t) \right| \right) . \tag{2.8}$$

therefore,

$$\|\underline{y}\| = \|\underline{Ax}\| \leq \|\underline{A}\| \, \|\underline{x}\|. \tag{2.9}$$

Thus the norms chosen are compatible, and a supplementary condition i.e. equation (2.9) can be included in (2.3).

Definition 8: An operator $\underline{\Gamma}$ mapping a normed space $\underline{X}$ into itself is said to be "bounded" if there exists a constant $M<\infty$ such that,

$$\|\underline{\Gamma}(\underline{x})\| \leq M \|\underline{x}\| \quad \forall \ \underline{x} \in \underline{X}. \tag{2.10}$$

Definition 9: A "Banach space" is a complete, normed-linear vector space. A space is complete if every Cauchy sequence from the space has its limit in the space.

Definition 10: A bounded operator $\underline{\Gamma}$, which maps a Banach space $\underline{X}$ into itself, is said to satisfy a "Lipschitz condition" if,

$$\|\underline{\Gamma}(\underline{x}) - \underline{\Gamma}(\underline{y})\| \leq K \|\underline{x} - \underline{y}\|. \tag{2.11}$$

$$\forall\ \underline{x}, \underline{y}, \in \underline{X}$$

where $o < K < oo$.

If $K < 1$ a contraction condition exists on $\underline{X}$ i.e. the "distance" between the transformed vectors is less than the "distance" between the original vectors.

A contraction is norm dependent, so that a transformation may be a contraction in one norm but not in another. However, because the spaces being considered are finite dimensional the equivalent norm theorem, Porter [16], will hold. Thus a contraction under one norm will also be a contraction under a different norm.

By virtue of (2.3(d)) it is clear that (2.11) demands $\underline{\Gamma}$ be continuous.

Necessary and sufficient conditions for $\underline{\Gamma}$ to be continuous are presented in section 2.3. A further property which is essential to the ensuing mathematical development is the concept of convexity.

Definition 11: Let $\underline{D} \subset \underline{X}$, where $\underline{X}$ is a Banach space. If $\underline{x}, \underline{y} \in \underline{D}$ then the "line segment" connecting $\underline{x}$ and $\underline{y}$ is a subset of $\underline{D}$ given by

$$\{ \underline{z} \in D: \underline{z} = \alpha \underline{x} + \beta \underline{y},\ \alpha \geq o,\ \beta \geq o,\ \alpha + \beta = 1 \}$$

which is a statement of convexity.

The spaces considered are Banach and are all locally convex about some point $\underline{x}_o \in \underline{D}$.

Let $\underline{x}$ and $\underline{y}$ be vectors contained in an open sphere $\underline{D}$ of radius $r > o$, centred at some point $\underline{x}_o \in \underline{D}$. Then

$$||\underline{x} - \underline{x}_o|| < r$$

and $$||\underline{y} - \underline{x}_o|| < r.$$

Estimating the "distance" from $\underline{x}_o$ to a point $\underline{z} = \alpha\underline{x} + \beta\underline{y}$ on the line segment connecting $\underline{x}$ and $\underline{y}$, results in

$$\begin{aligned} ||\underline{z} - \underline{x}_o|| &= ||(1 - \beta)\ \underline{x} + \beta\underline{y} - \underline{x}_o|| \\ &= ||(1 - \beta)(\underline{x} - \underline{x}_o) + \beta(\underline{y} - \underline{x}_o)|| \\ &\leq ||(1 - \beta)(\underline{x} - \underline{x}_o)|| + ||\beta(\underline{y} - \underline{x}_o)|| \\ &\leq (1 - \beta)\ ||\underline{x} - \underline{x}_o|| + \beta\,||\underline{y} - \underline{x}_o|| \\ &\leq (1 - \beta)\ r + \beta r = r. \end{aligned}$$

<u>Conclusion</u>: Any point on the line segment connecting $\underline{x}$ and $\underline{y}$ belongs to $\underline{D}$. Therefore, the whole line segment belongs to $\underline{D}$.

<u>Definition 12</u>: The operator $\underline{\Gamma}$ is said to be "convex" if it satisfies the following condition, that is,

$$\underline{\Gamma}\ (\underline{x} + \underline{y}) \leq \underline{\Gamma}\ (\underline{x}) + \underline{\Gamma}\ (\underline{y}).$$

clearly from defintions 7, 11 and 12, it is seen that the norm is a convex functional.

A further concept which will be utilized in the sequel is the one dimensional intermediate value theorem. The difference quotient of differential calculus is written as

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

and involves the values of the function for distinct values of x, whereas the derivative of the function at a point gives no information about the function at other points. Principally the difference quotient illustrates the properties of the function "in the large". The derivative illustrates a local property, that is, "in the small".

Later it will be necessary to derive global properties of the function i.e. properties over the whole of the subspace, from the local properties of the derivative. This will necessitate using the intermediate value theorem.

Intermediate Value Theorem: If $f(x)$ is continuous in a closed interval, $o \leq x \leq x + h$, and is differentiable at every point of the open interval, $o < x < x + h$, then there exists at least one value where $o < \lambda < 1$; such that

$$\frac{f(x + h) - f(x)}{h} = f'(\xi)$$

where $\xi = x + \lambda h.$

The functions to be considered in the succeeding sections are elements of the space of continuous functions from a domain of R^n into R^n. This space is called a "function space" and is written as $C((\underline{S}), \underline{R}^n)$, where $\underline{S} \subset \underline{R}^n$.

Let

$$\underline{f} \in C((\underline{S}), R^n)$$

and

$$\underline{x}(t) \in C([a, b], \underline{S})$$

and the norm of

$$\|\underline{f}\| = \sup_{\substack{t \in [a, b] \\ \underline{x} \in (\underline{S})}} \|\underline{f}\ (\underline{x}(t))\|.$$

Thus the function space is modified and may be written as

$$\underline{f}\ (\underline{x}(t)) \in C([a, b], (\underline{S}), \underline{R}^n)$$

where

$$C([a, b], (\underline{S}), \underline{R}^n) = \{\underline{f}(\underline{x}(t)):\ \underline{f} \text{ is continuous}$$

on a domain $\underline{S} \subset \underline{R}^n$ over $t \in [a, b]$, into $R^n\}$ [17].

2.2 The Frechet Differential

The differential has, by itself, influenced the development of mathematics more than any other single concept. It is one of the most important concepts used in non-linear systems analysis. A differential can play more than one major role significant in the theory and application of non-linear functional analysis, Nashed [18]. The role of major interest herein is that of the local approximation of non-linear operator by a linear operator.

Continuity of the operator is one of the prime pre-requisites for the application of the Banach contraction - mapping fixed point theorem, as has been stated in section 2.1.

In elementary calculus the concept of the partial derivative is used as a means of determining the derivative of a multi-variable function; each co-ordinate being considered separately.

Definition 14: If $\underline{x} = (x_1, x_2, ----, x_n)$ and $\underline{y} = (y_1, y_2, ----, y_n)$ are elements of a vector space $\underline{X}$, and $\underline{y}$ differs from $\underline{x}$ only in one co-ordinate that is $x_i = y_i$ for all $i \neq j$. The "partial derivative" of f with respect to the jth co-ordinate is written as

$$\frac{\partial f(\underline{x})}{\partial x_j} = \lim_{y_j \to x_j} \frac{f(\underline{y}) - f(\underline{x})}{y_j - x_j} . \tag{2.12}$$

Therefore, the total differential of f can be written as

$$df = \sum_{j=1}^{n} \frac{\partial f(\underline{x})}{\partial x_j} dx_j \tag{2.13}$$

It is well known that functions which exist in E^1 are continuous at a given point if the derivative exists at that point. Ortega [19] has demonstrated that a function in "n" (n>1) variables can have partial derivatives with respect to each of its variables and yet still be discontinuous. Clearly a stronger condition, than the existence of the partials, is required to imply continuity of the operator.

A stronger condition for continuity can be obtained via the directional derivative.

Definition 15: A function f is defined on an open set $\underline{S} \subset \underline{X}$ and $\underline{x} \in \underline{S}$; assume that $\underline{h}$ is a unit vector in $\underline{X}$. The "directional derivative" of f at the point $\underline{x}$ in the direction $\underline{h}$ is then

$$\lim_{\lambda \to o} \frac{f(\underline{x} + \lambda \underline{h}) - f(\underline{x})}{\lambda} \tag{2.14}$$

if the limit exists.

Since $\underline{h}$ is completely arbitrary, (2.14) must contain (2.12) as a special case, i.e. a specific direction is considered.

Definition 16: An "open set" $\underline{S} \subset \underline{X}$ has the following property:

For all $\underline{x} \in \underline{S}$ there exists an $r > o$ such that if a neighborhood $N(\underline{x},r)$ is erected about $\underline{x}$, and $\underline{y} \in N(\underline{x},r)$ then the distance between $\underline{x}$ and $\underline{y}$, $d(\underline{x},\underline{y})<r$.

Let $\varepsilon = r - d(\underline{x},\underline{y})$ then

$N(\underline{x},\varepsilon) \subset N(\underline{x},r)$.

Open sets are considered for the following reason. For all $\underline{x} \in \underline{S} \subset \underline{X}$ there will be a neighborhood about $\underline{x}$ completely contained in $\underline{S}$, i.e. $(N(\underline{x},\cdot)\subset\underline{S}$. Every point $\underline{y} \in N(\underline{x},\cdot)$ can be written in the following form: $\underline{y} = \underline{x} + \alpha\underline{h}$, where $\underline{h} \in \underline{X}$ and α is sufficiently small such that $\underline{x} + \alpha\underline{h} \in N(\underline{x},\cdot)$. Thus one sided or infinite derivatives can be neglected.

Further, it is to be noted that no restrictions are imposed upon $\underline{y}$ as to the way in which it is allowed to approach $\underline{x}$. Which implies that the neighborhood $N(\underline{x},\cdot)$ must be convex.

If a function has a directional derivative in every $\underline{h} \in \underline{X}$ then all the partials must exist. Yet the function may be discontinuous at the

given point [19]. Thus the existence of all the partials fails to imply continuity at the point. Clearly, the extension of a system from $n = 1$ to $n > 1$ cannot be obtained via the directional derivative.

A generalization of the directional derivative, introduced by Gateaux (1913) [19], is the differential. It does imply continuity and, at the same time, permits the extension of the principal theorems of derivatives from dimensions of $n = 1$ to $n > 1$.

<u>Definition 17</u>: A function $\underline{\Gamma}$ maps an open domain $\underline{S}$ of a vector space $\underline{X}$ into another vector space $\underline{Y}$, i.e.

$$\underline{\Gamma}: \ \underline{S} \subset \underline{X} \rightarrow \underline{Y} \ .$$

It has a "differential" at $\underline{x} \in \underline{S}$ if there exists another function $\underline{g}$ which satisfies the following conditions:

(1) $\underline{g}$ is a function of two variables, both of which have n dimensional bases. The differential of $\underline{\Gamma}$ is written as $\underline{g}(\underline{x};\underline{h})$. Where $\underline{x} \in \underline{S}$ is the point and $\underline{h} \in \underline{X}$ is the direction under consideration.

(2) $\underline{g}$ is homogeneous in $\underline{h}$ and of degree one.

(3) The differential must exist. It must be noted that no guarantee is given in advance that $\underline{\Gamma}$ has a differential. However, if $\underline{\Gamma}$ is suitably restricted - as will be shown later - then $\underline{g}$ can be shown to be unique. The differential of $\underline{\Gamma}$ at $\underline{x}$ in the direction $\underline{h}$ can also be written as

$$\underline{g}(\underline{x};\underline{h}) = \underline{\Gamma}'(\underline{x})\underline{h} \ , \tag{2.15}$$

where $\underline{\Gamma}'$ is the derivative of $\underline{\Gamma}$ at $\underline{x}$ and $\underline{\Gamma}'(\underline{x}) < \infty$.

<u>Definition 18</u>: The "Gateaux differential" of a function $\underline{\Gamma}$: $\underline{X} \to \underline{Y}$ is written as,

$$D\underline{\Gamma}(\underline{x})\underline{h} = \lim_{t \to o} \frac{\underline{\Gamma}(\underline{x} + t\underline{h}) - \underline{\Gamma}(\underline{x})}{t} = \underline{\Gamma}'(\underline{x})\underline{h}, \tag{2.16}$$

if the limit exists. Where $\underline{Y}$ is a Banach space, but $\underline{X}$ need not be. Clearly conditions (1) and (3) of definition 17 are satisfied. Proof of condition (2), definition 17, is now presented, that is,

$$D\underline{\Gamma}(\underline{x})(\lambda\underline{h}) = \lambda D\underline{\Gamma}(\underline{x})\underline{h}.$$

<u>Proof</u>:

For all $\lambda \varepsilon E'$, replace t by $t\lambda$ in (2.16) then

$$D\underline{\Gamma}(\underline{x})\underline{h} = \lim_{t \to o} \frac{\underline{\Gamma}(\underline{x} - t\lambda\underline{h}) - \underline{\Gamma}(\underline{x})}{t\lambda} = \frac{1}{\lambda} D\underline{\Gamma}(\underline{x})(\lambda\underline{h}).$$

<u>Conclusion</u>:

$$\lambda D\underline{\Gamma}(\underline{x})\underline{h} = D\underline{\Gamma}(\underline{x})(\lambda\underline{h}), \tag{2.17}$$

the homogeneity condition is satisfied.

Definition 18 satisfies all the conditions stated in definition 17, i.e. equation (2.16) is valid. Thus a function $\underline{\Gamma}$ has a G differential when (2.16) holds in every direction. Then $D\underline{\Gamma}(\underline{x})$ assigns to each vector $\underline{h} \varepsilon \underline{X}$ a vector $D\underline{\Gamma}(\underline{x})\underline{h} \varepsilon \underline{Y}$, therefore, $D\underline{\Gamma}(\underline{x})$ is a mapping.

It has been shown, by Ortega and Rheinbolt [19], that the G differential is not necessarily a linear operator. The G differential is lacking in the following desirable properties:

(1) it is not necessarily linear

(2) it has been shown by Nashed [18], that a G differential can exist and be a continuous linear operator while the function itself is discontinous.

Clearly then a stronger form of differential is required to overcome these deficiencies. The differential now to be considered is the F differential.

Definition 19: A mapping $\underline{\Gamma}$: $\underline{D} \subset \underline{X} \to \underline{Y}$ (where $\underline{D}$ is an open domain of $\underline{X}$) $\underline{X}$ and $\underline{Y}$ are normed real vector spaces; $\underline{\Gamma}$ is said to be "F differentiable" at $\underline{x} \in \underline{D}$ if the following representation holds for all $\underline{h} \in \underline{X} \ni \underline{x} + \underline{h} \in \underline{D}$,

$$\underline{\Gamma}(\underline{x} + \underline{h}) - \underline{\Gamma}(\underline{x}) = \underline{L}(\underline{x})\underline{h} + \underline{w}(\underline{x}; \underline{h}). \tag{2.18}$$

Where $\underline{L}(\underline{x})$ is a continuous linear operator and $\underline{L}(\underline{x})$: $\underline{D} \subset \underline{X} \to \underline{Y}$. Then $\underline{L}(\underline{x})\underline{h}$ is called the F differential of $\underline{\Gamma}$ at $\underline{x}$ and is written

$$\underline{L}(\underline{x})\underline{h} = \underline{\Gamma}'(\underline{x})\underline{h} \tag{2.19}$$

Thus $\underline{\Gamma}'(\underline{x}) \in \mathcal{L}[\underline{x}, \underline{y}]$, where $\mathcal{L}[\underline{x}, \underline{y}]$ is the space of linear operators which take a domain of $\underline{X}$ into $\underline{Y}$. Since the spaces $\underline{X}$ and $\underline{Y}$ are both normed real vector spaces, it will now be shown that if the F differential exists it is continuous.

Substitute (2.19) into (2.18) and take the norm of both sides,

$$\frac{\|\underline{\Gamma}(\underline{x} + \underline{h}) - \underline{\Gamma}(\underline{x}) - \underline{\Gamma}'(\underline{x})\underline{h}\|}{\|\underline{h}\|} = \frac{\|\underline{w}(\underline{x}; \underline{h})\|}{\|\underline{h}\|} . \tag{2.20}$$

It is to be noted here that $\underline{w}(\underline{x}; \underline{h})$ $\underline{Y}$ it is a mapping, dependent upon two n dimensional vectors and is called the remainder of the

differential. The map is defined for all sufficiently small vectors $\underline{h} \in \underline{X} (\underline{h} \neq \theta)$ such that

$$\lim_{\|\underline{h}\| \to o} \frac{\|\underline{w}(\underline{x};\ \underline{h})\|}{\|\underline{h}\|} = o. \tag{2.21}$$

The limit is obviously taken for $\|\underline{h}\| \neq o$, otherwise the quotient does not make sense.

Then

$$\lim_{\|\underline{h}\| \to o} \frac{\|\underline{\Gamma}(\underline{x} + \underline{h}) - \underline{\Gamma}(\underline{x}) - \underline{\Gamma}'(\underline{x})\underline{h}\|}{\|\underline{h}\|} = o \tag{2.22}$$

which is an alternative method of defining the F differential [20].

Because of (2.21) the F differential, $\underline{\Gamma}'(\underline{x})\underline{h}$, is continuous. The proof, part of which is reproduced from Vainberg [20] will now be presented.

<u>Proof</u>: $\underline{\Gamma}$: $\underline{D} \subset \underline{X} \to \underline{y}$ and if $\underline{\Gamma}$ is differentiable at some $\underline{x} \in \underline{D}$ then given $\varepsilon > o$ there exists $N(\underline{x}, \delta)$ such that if $\underline{x} + \underline{h} \in N(\underline{x}, \delta)$, then

$$\|\underline{\Gamma}(\underline{x} + \underline{h}) - \underline{\Gamma}(\underline{x}) - \underline{\Gamma}'(\underline{x})\underline{h}\| \leq \varepsilon \|\underline{h}\|. \tag{2.23}$$

Let

$$\|\underline{h}\| = \|\underline{y} - \underline{x}\| < \delta \tag{2.24}$$

where $\underline{y} \in N(\underline{x}, \delta)$,

now apply (2.3(d)) to the L.H.S of (2.23) thus

$$\begin{aligned} &\|\underline{\Gamma}(\underline{x} + \underline{h}) - \underline{\Gamma}(\underline{x})\| - \|\underline{\Gamma}'(\underline{x})\underline{h}\| \\ &\leq \|\underline{\Gamma}(\underline{x} + \underline{h}) - \underline{\Gamma}(\underline{x}) - \underline{\Gamma}'(\underline{x})\underline{h}\| . \end{aligned} \tag{2.25}$$

If (2.25) is now substituted into (2.23) then

$$\begin{aligned} \|\underline{\Gamma}(\underline{x} + \underline{h}) - \underline{\Gamma}(\underline{x})\| &\leq \varepsilon \|\underline{h}\| + \|\underline{\Gamma}'(\underline{x})\underline{h}\| \\ &\leq \varepsilon \|\underline{h}\| + \|\underline{\Gamma}'(\underline{x})\| \, \|\underline{h}\| . \end{aligned} \tag{2.26}$$

Now substitute (2.24) into (2.26) and write

$$\|\underline{\Gamma}(\underline{y}) - \underline{\Gamma}(\underline{x})\| \leq M\|\underline{y} - \underline{x}\| \tag{2.27}$$

where

$$M = \|\underline{\Gamma}'(\underline{x})\| + \varepsilon . \tag{2.28}$$

Since $\varepsilon > o$, let

$$\delta_1 = \min(\frac{\varepsilon}{M}; \ \delta) . \tag{2.29}$$

then

$$\begin{aligned} \|\underline{\Gamma}(\underline{y}) - \underline{\Gamma}(\underline{x})\| &\leq M\frac{\varepsilon}{M} \\ &\leq \varepsilon , \end{aligned} \tag{2.30}$$

when

$$\|\underline{y} - \underline{x}\| \leq \delta_1 . \tag{2.31}$$

<u>Conclusion:</u>

Since ε is arbitrary and positive this implies that $\underline{\Gamma}$ is continuous, and because of (2.3(d)) this implies that $\underline{\Gamma}'(\underline{x})\underline{h}$ is also continuous.

The F differential is utilized because it satisfies the following conditions:

(1) the F differential is defined in norm, thus it is a strong differential.

(2) if the F differential exists then

(a) the G differential exists

(b) $\underline{\Gamma}$ is continuous

(c) the derivative is unique

(d) the derivative is a continuous linear operator.

(3) by comparing (2.20) and (2.23) and observing the condition displayed in (2.24), then

$$\|\underline{h}\| \leq \delta \Rightarrow \|\underline{w}(\underline{x};\ \underline{h})\| \leq \varepsilon \|\underline{h}\|, \tag{2.32}$$

that is, $\underline{\Gamma}$ is uniformly differentiable on $\underline{D}$ with respect to all $\underline{x} \in \underline{D}$.

(4) $\underline{\Gamma}'(\underline{x})$ is continuous in $\underline{D}$ iff $\underline{\Gamma}$ has a locally uniform differential - condition (3) exists - and if $\underline{\Gamma}'(\underline{x})$ is locally bounded on $\underline{D}$.

Locally bounded means every $\underline{x} \in \underline{D}$ has a neighborhood $N(\underline{x}, \cdot)$ in which $\|\underline{\Gamma}'(\underline{x})\|$ is bounded, Vainberg [20].

Bounded in norm is an immediate consequence of continuity; $\underline{\Gamma}'(\underline{x})$ is uniformly continuous iff $\underline{\Gamma}(\underline{x})$ is uniformly differentiable.

(5) if $\underline{\Gamma}'(\underline{x})$ is uniformly continuous in an open-bounded convex domain $\underline{D}$, then the F differential is uniform in $\underline{D}$.

Details of the proofs for conditions (4) and (5) are to be found in [18]-[21].

By observing condition (2) proof of existence of the F differential is a prime criterion. A proof of this is offered in section 2.3.

2.3 The Existence of the F Derivative

In the latter part of section 2.2 reasons for utilizing the F derivative were stated. One condition which emerged as being of paramount importance was that the F differential must exist.

In this section sufficient conditions for the proof of existence are presented. The presentation is made in the following manner:

(1) if the F derivative exists, then proofs are presented which will show that,

(a) the partial derivatives must exist and be continuous,

(b) the Lipschitz condition is satisfied for a multi-variable system.

(2) if the partial derivatives exist and are continuous, then the F differential exists; sufficient conditions only.

(3) from the existence of the F differential it will be shown that the F derivative exists and is a linear operator.

(4) an argument is presented which substantiates the unique representation of the F derivative, - if it exists and is a linear operator - by a Jacobian matrix.

For convenience this representation will be used early in the section, but will not be justified until later.

Figure 2 is instructive in the development of the F derivative from the G differential.

Suppose that the F differential does exist at a point $\underline{x} \in \underline{D}$ in the direction $\underline{h} \in \underline{X}$. Previously it was stated, equation (2.15), that the F differential could be represented by $\underline{g}(\underline{x}; \underline{h})$ where

$$\underline{g} = (g_1, g_2, \text{----}, g_n) \qquad (2.33)$$

and

$$\underline{h} = (h_1, h_2, ----, h_n) \tag{2.34}$$

Proof of existence of the partial derivatives, based on the hypothesis that the F differential exists and can be represented as above, will now be presented.

Proof:

A vector $\underline{u}$ is defined as $\underline{u} = (u_1, u_2, ----, u_n)$ where u_1, u_2 etc. are the unit vectors which form the basis of the vector space $\underline{X} \subset R^n$. A component of the vector $\underline{u}$ i.e. u_j, is defined as $u_j = (o, o, ---, o, 1, o, ----, o)$; a 1 in the jth position. Thus $\underline{h}$ can be uniquely written as

$$\underline{h} = h_1u_1 + h_2u_2 + ---- + h_nu_n \tag{2.35}$$

$$= \sum_{j=1}^{n} h_ju_j , \tag{2.36}$$

then

$$g_i(\underline{x};\ \underline{h}) = g_i(\underline{x}; \sum_{j=1}^{n} h_ju_j) . \tag{2.37}$$

Since $\underline{g}(\underline{x};\ \underline{h})$ is homogeneous in $\underline{h}$ of degree one, equation (2.37) can be written

$$g_i(\underline{x}; \sum_{j=1}^{n} h_ju_j) = g_i(\underline{x};\ h_1u_1) + g_i(\underline{x};\ h_2u_2) + ---$$

$$--- + --- + g_i(\underline{x};\ h_nu_n) \tag{2.38}$$

$$= g_i(\underline{x};\ u_1)h_1 + g_i(\underline{x};\ u_2)h_2 + \text{---}$$

$$\text{---} + \text{---} + g_i(\underline{x};\ u_n)h_n \tag{2.39}$$

$$= a_{i1}(\underline{x})h_1 + a_{i2}(\underline{x})h_2 + \text{---}$$

$$\text{---} + \text{---} + a_{in}(\underline{x})h_n \tag{2.40}$$

$$= \sum_{j=1}^{n} a_{ij}(\underline{x})h_j \quad , \tag{2.41}$$

where $$a_{ij}(\underline{x})h_j = g_i(\underline{x};\ u_j)h_j \quad . \tag{2.42}$$

From equation (2.24) $\underline{h} = \underline{y} - \underline{x}$ therefore,

$$h_j = y_j - x_j \quad \forall\ j = 1, 2, \text{---}, n. \tag{2.43}$$

If $\underline{x}$ is allowed to vary in one direction only, at a time, then

$$\underline{y} = \underline{x} + \lambda u_j \tag{2.44}$$

where $o < |\lambda| < \delta$ then

$$o < ||\underline{y} - \underline{x}|| = ||\lambda|| \quad . \tag{2.45}$$

Thus

$$y_j - x_j = \lambda = h_j \tag{2.46}$$

and $y_k - x_k = o$ for $k \neq j$, therefore $h_k = o$.

Substituting h_j and h_k into equation (2.41)

$$g_i(\underline{x};\ \underline{y} - \underline{x}) = a_{ij}(\underline{x})\lambda \quad , \tag{2.47}$$

since all the other terms of the summation are zero, by virtue of the fact that $\underline{x}$ differs from $\underline{y}$ only in the jth co-ordinate. If equation (2.15) is now substituted into (2.23) and only the ith component of $\underline{\Gamma}$ is considered then

$$||\Gamma_i(\underline{x} + \lambda \underline{u}_j) - \Gamma_i(\underline{x}) - g_i(\underline{x}; \underline{y} - \underline{x})|| \leq \varepsilon ||h_j|| . \qquad (2.48)$$

Now substitute (2.47) and $h_j = \lambda$ into (2.48).

Thus

$$\left|\left| \frac{\Gamma_i(\underline{x} + \lambda \underline{u}_j) - \Gamma_i(\underline{x})}{\lambda} - a_{ij}(\underline{x}) \right|\right| \leq \varepsilon \qquad (2.49)$$

and

$$\lim_{\lambda \to o} \frac{\Gamma_i(\underline{x} + \lambda \underline{u}_j) - \Gamma_i(\underline{x})}{\lambda} = \frac{\partial \Gamma_i(\underline{x})}{\partial x_j} \qquad (2.50)$$

Because ε is arbitrary and positive this implies

$$a_{ij}(\underline{x}) = \frac{\partial \Gamma_i(\underline{x})}{\partial x_j} \qquad (2.51)$$

and also that $a_{ij}(\underline{x})$ exists; therefore,

$$a_{ij}(\underline{x}) = g_i(\underline{x}; u_j) = \frac{\partial \Gamma_i(\underline{x})}{\partial x_j} . \qquad (2.52)$$

Conclusion:

(1) $g_i(\underline{x}; u_j)$ can be uniquely represented by $a_{ij}(\underline{x})$.

(2) Because of the continuity condition (equation (2.49)) the partial derivatives exist.

Therefore, if the F differential exists then all the partial derivatives must exist.

Substitute equation (2.52) into (2.41) and write,

$$g_i(\underline{x};\ \underline{h}) = \sum_{j=1}^{n} \frac{\partial \Gamma_i(\underline{x})}{\partial x_j} h_j . \tag{2.53}$$

Replace h_j by equation (2.43) then

$$g_i(\underline{x};\ \underline{h}) = \sum_{j=1}^{n} \frac{\partial \Gamma_i(\underline{x})}{\partial x_j} (y_j - x_j) . \tag{2.54}$$

Writing equation (2.48) in a more general form gives,

$$\|\Gamma_i(\underline{x} + \underline{h}) - \Gamma_i(\underline{x}) - g_i(\underline{x};\ \underline{h})\| \leq \varepsilon \|\underline{h}\| \tag{2.55}$$

which can be rewritten as

$$\|\Gamma_i(\underline{x} + \underline{h}) - \Gamma_i(\underline{x})\| \leq \varepsilon \|\underline{h}\| + \|g_i(\underline{x};\ \underline{h})\| . \tag{2.56}$$

Taking the norm of (2.54)

$$\|g_i(\underline{x};\ \underline{h})\| = \left\|\sum_{j=1}^{n} \frac{\partial \Gamma_i(\underline{x})}{\partial x_j} (y_j - x_j)\right\| .$$

$$\leq \left\|\sum_{j=1}^{n} \frac{\partial \Gamma_i(\underline{x})}{\partial x_j}\right\| \|\underline{y} - \underline{x}\|$$

$$\leq \sum_{j=1}^{n} \left\|\frac{\partial \Gamma_i(\underline{x})}{\partial x_j}\right\| \|\underline{y} - \underline{x}\| , \tag{2.57}$$

but $\underline{h} = \underline{y} - \underline{x}$ therefore,

$$\|\underline{h}\| = \|\underline{y} - \underline{x}\| . \tag{2.58}$$

Therefore if (2.57) and (2.58) are now substituted into the R.H.S of (2.56) then it may be rewritten as

$$\|\Gamma_i(\underline{x} + \underline{h}) - \Gamma_i(\underline{x})\| \leq \|\underline{y} - \underline{x}\| \left(\varepsilon + \sum \left\| \frac{\partial \Gamma_i(\underline{x})}{\partial x_j} \right\| \right) . \tag{2.59}$$

It was stated earlier that $\underline{\Gamma}'(\underline{x}) < \infty$ (by hypothesis) therefore, each component must also be bounded, that is

$$\Gamma'_i(\underline{x}) < \infty \qquad \forall\ i = 1, 2, ----, n \tag{2.60}$$

from which it follows that

$$\sum_{j=1}^{n} \left\| \frac{\partial \Gamma_i(\underline{x})}{\partial x_j} \right\| < \infty \quad . \tag{2.61}$$

Now let

$$M_i = \varepsilon + \sum_{j=1}^{n} \left\| \frac{\partial \Gamma_i(\underline{x})}{\partial x_j} \right\| \tag{2.62}$$

then

$$\|\Gamma_i(\underline{x} + \underline{h}) - \Gamma_i(\underline{x})\| \leq M_i \|\underline{y} - \underline{x}\| \ ; \tag{2.63}$$

thus the Lipschitz condition is satisfied on a component basis, because of the initial hypothesis that the F differential exists. Proof that the results stated in (2.63) can be extended to include all i, where $i = 1, 2, ---, n$ and be written as

$$\|\underline{\Gamma}(\underline{x} + \underline{h}) - \underline{\Gamma}(\underline{x})\| \leq M \|\underline{y} - \underline{x}\| \tag{2.64}$$

where

$$M = \sup_i \{M_i\} , \tag{2.65}$$

will now be presented. The proof is developed via the one dimensional intermediate value theorem, and utilizes the principle of convexity in Banach spaces.

Proof:

Consider a sphere $\underline{D} \subset \underline{X}$ of radius $r > o$, centred at some point $\underline{x}_o \in \underline{D}$, let $\underline{x}$, $y \in \underline{D}$ and let $\underline{\Gamma}$: $\underline{D} \subset \underline{X} \to \underline{X}$. A line segment is considered as connecting the two points $\underline{x}$ and $\underline{y}$; a point $\underline{z}$, which is on the line segment, is defined in terms of $\underline{x}$ and $\underline{y}$ by

$$\underline{z} = \underline{x} + \lambda(\underline{y} - \underline{x}), \quad o \leq \lambda \leq 1. \tag{2.66}$$

Keeping λ fixed and introducing a new function which depends only upon λ, i.e. $\underline{\gamma}(\lambda)$, which is the "distance" between the images of $\underline{x}$ and $\underline{x} + \lambda(\underline{y} - \underline{x})$ under $\underline{\Gamma}$, then

$$\underline{\gamma}(\lambda) = \underline{\Gamma}(\underline{x} + \lambda(\underline{y} - \underline{x})) - \underline{\Gamma}(\underline{x}) . \tag{2.67}$$

If now $\lambda = 1$, then

$$\underline{\gamma}(1) = \underline{\Gamma}(\underline{y}) - \underline{\Gamma}(\underline{x}) \tag{2.68}$$

and if $\lambda = o$ then

$$\underline{\gamma}(o) = \underline{\theta} \text{ (null vector)}, \tag{2.69}$$

therefore,

$$\underline{\gamma}(1) - \underline{\gamma}(o) = \underline{\Gamma}(\underline{y}) - \underline{\Gamma}(\underline{x}) \ . \tag{2.70}$$

Consider now the ith component of $\underline{\gamma}$, that is γ_i for i = 1, 2, ---, n, partition the interval [o, 1], $o = \lambda_o < \lambda_1 < --- < \lambda_n = 1$, and let λ_g and λ_k be two different but fixed points in the partition, therefore,

$$\begin{aligned}
&\gamma_i(\lambda_g) - \gamma_i(\lambda_k) \\
&= \Gamma_i(\underline{x} + \lambda_g(\underline{y} - \underline{x})) - \Gamma_i(\underline{x}) \\
&\quad - \Gamma_i(\underline{x} + \lambda_k(\underline{y} - \underline{x})) + \Gamma_i(\underline{x}) \\
&= \Gamma_i(\underline{x} + \lambda_g(\underline{y} - \underline{x})) - \Gamma_i(\underline{x} + \lambda_k(\underline{y} - \underline{x})) .
\end{aligned} \tag{2.71}$$

Now, by considering each component of $\underline{x}$ and $\underline{y}$ e.g. x_j, y_j typically, equation (2.71) can be rewritten as

$$\begin{aligned}
&\Gamma_i(x_j + \lambda_g(y_j - x_j)) - \Gamma_i(x_j + \lambda_k(y_j - x_j)) \\
&= \frac{\partial\Gamma_i(\cdot)}{\partial x_j}(\lambda_g - \lambda_k)(y_j - x_j) \ ,
\end{aligned} \tag{2.72}$$

where

$$\frac{\partial\Gamma_i(\cdot)}{\partial x_j} = \frac{\partial\Gamma_i(x_j + \theta(y_j - x_j))}{\partial x_j}$$

and

$$\lambda_g < \theta < \lambda_k;$$

$\dfrac{\partial\Gamma_i(\cdot)}{\partial x_j}$ is the mean slope in the interval (λ_g, λ_k). The material which was developed earlier and culminated in equation (2.53) is now utilized, that is, we can rewrite (2.71) as

$$\gamma_i(\lambda_g) - \gamma_i(\lambda_k) = \sum_{j=1}^{n} \frac{\partial\Gamma_i(\cdot)}{\partial x_j} (\lambda_g - \lambda_k)(y_j - x_j)$$

for all i = 1, 2, ---, n.

But $$\frac{d\gamma_i(\cdot)}{d\lambda_g} = \frac{\gamma_i(\lambda_g) - \gamma_i(\lambda_k)}{\lambda_g - \lambda_k} \tag{2.73}$$

where $\dfrac{d\gamma_i(\cdot)}{d\lambda_g}$ is the mean slope in the interval (λ_g, λ_k).

Therefore,

$$\frac{\gamma_i(\lambda_g) - \gamma_i(\lambda_k)}{\lambda_g - \lambda_k} = \sum_{j=1}^{n} \frac{\partial\Gamma_i(\cdot)}{\partial x_j} (y_j - x_j) \tag{2.74}$$

$$= \frac{d\gamma_i(\cdot)}{d\lambda_g}$$

and

$$\frac{d\underline{\gamma}(\cdot)}{d\lambda} = \begin{bmatrix} \frac{d\gamma_i(\cdot)}{d\lambda} \\ | \\ | \\ | \\ | \\ | \\ \frac{d\gamma_n(\cdot)}{d\lambda} \end{bmatrix} \quad \text{a column vector}$$

The general form of equation (2.74) is

$$\frac{d\underline{\gamma}(\cdot)}{d\lambda} = \left[\frac{\partial\Gamma_i(\cdot)}{\partial x_j}\right] (\underline{y} - \underline{x}) \tag{2.75}$$

where $\left[\frac{\partial\Gamma_i(\cdot)}{\partial x_j}\right]$ is the Jacobian matrix of $\underline{\Gamma}'(\cdot)$ [18], and is written

$$\underline{\Gamma}'(\cdot) = \begin{bmatrix} \frac{\partial\Gamma_1(\cdot)}{\partial x_1} & \frac{\partial\Gamma_1(\cdot)}{\partial x_2} & --- & \frac{\partial\Gamma_1(\cdot)}{\partial x_n} \\ | & & & | \\ | & & & | \\ | & & & \\ \frac{\partial\Gamma_n(\cdot)}{\partial x_1} & ----- & & \frac{\partial\Gamma_n(\cdot)}{\partial x_n} \end{bmatrix} \tag{2.76}$$

Writing $\frac{\partial\Gamma_i(\cdot)}{\partial x_j}$ in the form $a_{ij}(\cdot)$, equation (2.51), then (2.76) can be rewritten as

$$\underline{\Gamma}' = \begin{bmatrix} a_{11}(\cdot) & a_{12}(\cdot) & --- & a_{1n}(\cdot) \\ | & & & | \\ | & & & | \\ | & & & | \\ a_{n1}(\cdot) & & & a_{nn}(\cdot) \end{bmatrix} . \tag{2.77}$$

The elements of the Jacobian matrix are the first partial derivatives of the vector $\underline{x}$, with respect to its components. They were obtained via the one dimensional intermediate value theorem.

Now extend the region between λ_g and λ_k to $\lambda_g = 1$ and $\lambda_k = o$, then equation (2.74) can be rewritten as

$$\frac{\gamma_i(1) - \gamma_i(o)}{1 - o} = \frac{d\gamma_i(\bar{\lambda}_k)}{d\lambda_k} \tag{2.78}$$

where $\bar{\lambda}_k$ is an intermediate value of λ_k in the interval [0, 1]. But $\gamma_i(o) = o$ (equation (2.69)) therefore,

$$\gamma_i(1) = \frac{d\gamma_i(\bar{\lambda}_k)}{d\lambda_k} \tag{2.79}$$

and from equation (2.70)

$$\gamma_i(1) - \gamma_i(o) = \Gamma_i(\underline{y}) - \Gamma_i(\underline{x}) . \tag{2.80}$$

By virtue of (2.75) we can write

$$\begin{aligned} \frac{d_{\underline{\gamma}}(\cdot)}{d\lambda} &= \underline{\Gamma}(\underline{y}) - \underline{\Gamma}(\underline{x}) \\ &= \left[\frac{\partial\Gamma_i(\cdot)}{\partial x_j}\right](\underline{y} - \underline{x}) ; \end{aligned} \tag{2.81}$$

and taking the norm of (2.81)

$$\|\underline{\Gamma}(\underline{y}) - \underline{\Gamma}(\underline{x})\| = \left\| \left[\frac{\partial \Gamma_i(\cdot)}{\partial x_j}\right](\underline{y} - \underline{x}) \right\|$$

$$\leq \left\| \left[\frac{\partial \Gamma_i(\cdot)}{\partial x_j}\right] \right\| \, \|\underline{y} - \underline{x}\| \tag{2.82}$$

which is valid for all $\underline{x}$ and $\underline{y}$ in the neighborhood of $\underline{x}_o$ [22].

A comparison of (2.64) and (2.82) reveals that

$$M = \varepsilon + \sup_i \sum_{j=1}^{n} \left\| \frac{\partial \Gamma_i(\underline{x})}{\partial x_j} \right\| \tag{2.83}$$

and since ε is positive and arbitrary the equation (2.82) can be rewritten as

$$\|\underline{\Gamma}(\underline{y}) - \underline{\Gamma}(\underline{x})\| \leq M \|\underline{y} - \underline{x}\|. \tag{2.84}$$

Conclusion:

The Lipschitz condition is satisfied for multivariable systems.

The concept of the one dimensional intermediate value theorem has been used to extend the "local" properties of the F derivative to the "global" properties of the quotient, as discussed in section 2.1.

It was stated earlier in this section that sufficient conditions only exist in the proof of the F differential. These conditions will now be developed.

Proof:

Suppose that the partial derivatives do exist and are continuous at a point $\underline{x}$, then there exists an $\varepsilon > o$ such that if $\underline{y} \varepsilon N(\underline{x}, \delta/2)$ then this implies, Luenberger [31]

$$\left\| \frac{\partial \Gamma_i(\underline{y})}{\partial x_j} - \frac{\partial \Gamma_i(\underline{x})}{\partial x_j} \right\| < \frac{\varepsilon}{n} , \tag{2.85}$$

where j and i = 1,2,---, n. Because different neighborhoods exist for each component of $\underline{x}$ let

$$\delta' = \inf(\delta_1, \delta_2, ---, \delta_n).$$

Now let $\underline{N}_k = \sum_{j=1}^{k} h_j u_j$ such that

$N_o = \theta$ (null vector), and further $\underline{N}_k$ satisfies the following recurrence relationship

$$\underline{N}_k = \underline{N}_{k-1} + \underline{h}_k u_k , \tag{2.86}$$

therefore,

$$\| \underline{N}_k \| \leq \| \underline{h} \| \quad \text{for all k.} \tag{2.87}$$

But

$$\begin{aligned} \underline{y} &= \underline{x} + \underline{hu} \\ \| \underline{y} - \underline{x} \| &= \| \underline{h} \| = \frac{\delta'}{2} \end{aligned} \tag{2.88}$$

thus

$$\Gamma_i(\underline{y}) - \Gamma_i(\underline{x}) = \Gamma_i(\underline{x} + \underline{hu}) - \Gamma_i(\underline{x}), \tag{2.89}$$

now because of $\underline{N}_k = \sum_{j=1}^{k} \underline{h}_j u_j$ we can write (2.89) as follows

$$\Gamma_i(\underline{x} + \underline{hu}) - \Gamma_i(\underline{x}) = \sum_{k=1}^{n} \Gamma_i(\underline{x} + \underline{N}_k) - \Gamma_i(\underline{x} + \underline{N}_{k-1}) \tag{2.90}$$

and by substituting (2.86) into (2.90) the general term of (2.90) may be written as

$$\Gamma_i(\underline{x} + \underline{N}_k) - \Gamma_i(\underline{x} + \underline{N}_{k-1}) = \Gamma_i(\underline{x} + \underline{N}_{k-1} + \underline{h}_k u_k) - \Gamma_i(\underline{x} + \underline{N}_{k-1}) . \tag{2.91}$$

Which is the difference of one component of the operator $\underline{\Gamma}$ at two points, the two points differ only in the k^{th} co-ordinate in the neighborhood $N(\underline{x}, \delta'/2)$.

A new function is introduced at this point

$$P(\underline{h}_k u_k) = \Gamma_i(\underline{x} + \underline{N}_{k-1} + \underline{h}_k u_k) \tag{2.92}$$

$$P(\underline{\theta}) = \Gamma_i(\underline{x} + \underline{N}_{k-1}) \tag{2.93}$$

and the intermediate value theorem is utilized again to write

$$\frac{P(\underline{h}_k u_k) - P(\underline{\theta})}{\underline{h}_k - \underline{\theta}} = P'(\underline{\bar{h}}_k) \tag{2.94}$$

where

$$\underline{\theta} < \underline{\bar{h}}_k u_k \leq \underline{h}_k u_k . \tag{2.95}$$

Then

$$\frac{P(\bar{\underline{h}}_k u_k + \underline{\gamma} u_k) - P(\bar{\underline{h}}_k u_k)}{\underline{\gamma}}$$

$$= \frac{\Gamma_i(\underline{x} + \underline{N}_{k-1} + \bar{\underline{h}}_k u_k + \underline{\gamma} u_k) - \Gamma_i(\underline{x} + \underline{N}_{k-1} + \bar{\underline{h}}_k u_k)}{\underline{\gamma}} \tag{2.96}$$

taking (2.96) to the limit, $\underline{\gamma} \to \underline{\theta}$ then

$$\lim_{\underline{\gamma} \to \underline{\theta}} \frac{P(\bar{\underline{h}}_k u_k + \underline{\gamma} u_k) - P(\bar{\underline{h}}_k u_k)}{\underline{\gamma}}$$

$$= \frac{\partial \Gamma_i(\underline{x} + \underline{N}_{k-1} + \bar{\underline{h}}_k u_k)}{\partial x_k} . \tag{2.97}$$

Now rewrite (2.94) as follows

$$P(\underline{h}_k u_k) - P(\underline{\theta}) = \underline{h}_k P'(\bar{\underline{h}}_k u_k) \tag{2.98}$$

and substitute equations (2.92), (2.93) and (2.97) into (2.98) which results in

$$\Gamma_i(\underline{x} + \underline{N}_{k-1} + \underline{h}_k u_k) - \Gamma_i(\underline{x} + \underline{N}_{k-1})$$

$$= \underline{h}_k \frac{\partial \Gamma_i(\underline{x} + \underline{N}_{k-1} + \bar{\underline{h}}_k u_k)}{\partial x_k} . \tag{2.99}$$

Substitute (2.99) into (2.90) to obtain

$$\Gamma_i(\underline{x} + \underline{h}u) - \Gamma_i(\underline{x}) = \sum_{k=1}^{n} \underline{h}_k \frac{\partial \Gamma_i(\underline{x} + \underline{N}_{k-1} + \bar{\underline{h}}_k u_k)}{\partial x_k} u_k$$

$$= \sum_{k=1}^{n} \underline{h}_k \left(\frac{\partial \Gamma_i(\underline{x} + \underline{N}_{k-1} + \bar{\underline{h}}_k u_k)}{\partial x_k} - \frac{\partial \Gamma_i(\underline{x})}{\partial x_k} + \frac{\partial \Gamma_i(\underline{x})}{\partial x_k} \right) u_k$$

$$= \sum_{k=1}^{n} \underline{h}_k \frac{\partial \Gamma_i(\underline{x})}{\partial x_k} u_k + \sum_{k=1}^{n} \underline{h}_k \left(\frac{\partial \Gamma_i(\underline{x} + \underline{N}_{k-1} + \bar{\underline{h}}_k u_k}{\partial x_k} - \frac{\partial \Gamma_i(\underline{x})}{\partial x_k} \right) u_k \quad (2.100)$$

and from equation (2.53) it is seen that

$$g_i(\underline{x};\ \underline{h}) = \sum_{k=1}^{n} \underline{h}_k \frac{\partial \Gamma_i(\underline{x})}{\partial x_k} u_k. \quad (2.101)$$

The remainder of the terms in the R.H.S. of equation (2.100) are now considered, that is

$$\sum_{k=1}^{n} \underline{h}_k u_k \left(\frac{\partial \Gamma_i(\underline{x} + \underline{N}_{k-1} + \bar{\underline{h}}_k u_k)}{\partial x_k} - \frac{\partial \Gamma_i(\underline{x})}{\partial x_k} \right). \quad (2.102)$$

By combining (2.87) and (2.88) it is seen that

$$\|\underline{N}_{k-1} + \bar{\underline{h}}_k u_k\| \le \|\underline{h}\| + \|\underline{hu}\|$$
$$\le 2\|\underline{h}\| < \delta' . \tag{2.103}$$

Therefore, the combined effect of (2.85), (2.88) and (2.103) on (2.102) will cause this to reduce to $\varepsilon(\underline{h})$. Thus (2.100) can be rewitten as

$$\Gamma_i(\underline{x} + \underline{hu}) - \Gamma_i(\underline{x}) = g_i(\underline{x};\ \underline{h}) + \varepsilon(\underline{h}) \tag{2.104}$$

but since $\underline{x} + \underline{hu} = \underline{y}$ then (2.104) may be rewritten as

$$\Gamma_i(\underline{y}) = \Gamma_i(\underline{x}) = g_i(\underline{x};\ \underline{h}) + \varepsilon(\underline{h}) . \tag{2.105}$$

Therefore,

$$\|\Gamma_i(\underline{y}) - \Gamma_i(\underline{x}) - g_i(\underline{x};\ \underline{h})\| \le \varepsilon\|\underline{y} - \underline{x}\| \tag{2.106}$$

and since y is arbitrary in $N(\underline{x}, \delta'/2)$ the differential $g_i(\underline{x};\ \underline{h})$ exists.

Conclusion:

If the partial derivatives exist and are continuous then the F differential exists; sufficient conditions only.

Now that sufficient conditions have been established to prove the existence of the F differential it only remains to show that the F derivative exists.

In previous sections it was stated that the F differential was homogeneous in $\underline{h}$ of degree one. The proof was not presented at that time, this will now be corrected. Also proof that the F differential is linear will be offered. Sections of this proof are to be found in Vainberg [20].

Proof:

Suppose that Γ: $\underline{D} \subset \underline{X} \rightarrow \underline{Y}$ has at a point $\underline{x} \in \underline{D}$ an F differential then there exists an $\varepsilon > o$ such that if $\underline{y} \in N(\underline{x}, \delta)$ this implies that the F differential is linear.

$$\underline{\Gamma}(\underline{x} + \underline{h}) - \underline{\Gamma}(\underline{x}) = \underline{g}(\underline{x}; \underline{h}) + \underline{\alpha}_1 \tag{2.107}$$

$$\underline{\Gamma}(\underline{x} + \underline{m}) - \underline{\Gamma}(\underline{x}) = \underline{g}(\underline{x}; \underline{m}) + \underline{\alpha}_2 \tag{2.108}$$

and

$$\underline{\Gamma}(\underline{x} + \underline{m} + \underline{h}) - \underline{\Gamma}(\underline{x}) = \underline{g}(\underline{x}; \underline{m} + \underline{h}) + \underline{\alpha}_3 \tag{2.109}$$

where

$$\|\underline{\alpha}_i\| < \frac{\varepsilon}{4} \qquad i = 1, 2, 3, 4.$$

Combine equations (2.107), (2.108) and (2.109) and take the norm of both sides, then

$$\left\|\underline{g}(\underline{x};\ \underline{m}+\underline{h}) - \underline{g}(\underline{x};\ \underline{m}) - \underline{g}(\underline{x};\ \underline{h})\right\| \leq$$

$$\left\|\underline{\Gamma}(\underline{x}+\underline{m}+\underline{h}) - \underline{\Gamma}(\underline{x}+\underline{m}) - \underline{\Gamma}(\underline{x}+\underline{h}) - \underline{\Gamma}(\underline{x})\right\| + \frac{3\varepsilon}{4} \qquad (2.110)$$

but

$$\underline{\Gamma}(\underline{x}+\underline{m}+\underline{h}) - \underline{\Gamma}(\underline{x}+\underline{m}) = \underline{g}(\underline{x}+\underline{m};\ \underline{h}) + \underline{\alpha}_4. \qquad (2.111)$$

Now substitute equations (2.107)and (2.111) into (2.110) and write

$$\left\|\underline{g}(\underline{x};\ \underline{m}+\underline{h}) - \underline{g}(\underline{x};\ \underline{m}) - \underline{g}(\underline{x};\ \underline{h})\right\| \leq$$
$$\left\|\underline{g}(\underline{x}+\underline{m};\ \underline{h}) - \underline{g}(\underline{x};\ \underline{h})\right\| + \frac{3\varepsilon}{4}\ . \qquad (2.112)$$

By generalizing equation (2.55) it is seen that $\underline{g}(\underline{x};\ \cdot)$ exists and is continuous, therefore, by suitable choice of δ

$$\left\|\underline{g}(\underline{x}+\underline{m};\ \underline{h}) - \underline{g}(\underline{x};\ \underline{h})\right\| < \frac{\varepsilon}{4} \qquad (2.113)$$

and if (2.113) is substituted into (2.112) then

$$\left\|\underline{g}(\underline{x};\ \underline{m}+\underline{h}) - \underline{g}(\underline{x};\ \underline{m}) - \underline{g}(\underline{x};\ \underline{h})\right\| \leq \varepsilon\ . \qquad (2.114)$$

Since ε is arbitrary and positive then

$$\underline{g}(\underline{x};\ \underline{m}+\underline{h}) = \underline{g}(\underline{x};\ \underline{m}) + \underline{g}(\underline{x};\ \underline{h}), \qquad (2.115)$$

Conclusion:

The additivity property of $\underline{g}(;)$ is satisfied. It will be remembered that the homogeneity condition was satisfied in equation (2.17). Therefore, it can be concluded that the F differential is linear in the variable $\underline{h}$, and thus exists.

It can be seen from equation (2.52) that $\underline{g}(\underline{x};\ \underline{h}) = \underline{\Gamma}'(\underline{x})\underline{h}$. The operator which assigns $\underline{\Gamma}'(\underline{x})$ to $\underline{x}$ is called the F derivative of $\underline{\Gamma}$. The derivative has its domain in the same space as the original operator, also it has the same range space as the original operator.

Matrices are closely related to linear transformations; this fact is well known. Consider the following system of "n" linear equations

$$\begin{array}{lcl} a_{11}h_1 + a_{12}h_2 + - - - - & + a_{1n}h_n & = Z_1 \\ a_{21}h_1 + a_{22}h_2 + - - - - & + a_{2n}h_n & = Z_2 \\ \quad | & \quad | & \ \ | \\ \quad | & \quad | & \ \ | \\ \quad | & \quad | & \ \ | \\ \quad | & \quad | & \ \ | \\ a_{n1}h_1 + - - - - - & + a_{nn}h_n & = Z_n \end{array} \tag{2.116}$$

it is equally well known that this set of "n" linear equations assigns a unique set of "n" variables $Z_1, Z_2, - - -, Z_n$ to every set of $h_1, h_2, - - -, h_n$ variables. This assignment is called a linear transformation and is characterized by an array of n^2 coefficients, which is a matrix. If the general row of (2.116) is compared to the general row from (2.53) it will be seen that they have the same form, providing (2.53) represents a linear equation. The proof, equations (2.107) to (2.115), has shown $\underline{\Gamma}'(\underline{x})$ to be a linear

operator. Its components can be shown to be the same. Thus, the comparison reveals equivalency. The justification for writing $\underline{\Gamma}'(\underline{x})$ as a Jacobian matrix was not revealed at the time that (2.76) was presented. In the light of the proof mentioned, i.e. the last proof in this section, and the material which followed, the representation is justified.

Conclusion:

If $\underline{\Gamma}'(\underline{x})$ exists and is a linear operator it may be uniquely represented by a Jacobian matrix, Nashed [18].

2.4 The Banach Contraction Mapping-Fixed Point Principle

It was stated earlier in this chapter that the Banach contraction mapping-fixed point principle would be used as a means of obtaining a solution to the non-linear mapping problem. For brevity it will be called the "fixed point principle".

A statement exists which asserts that under certain conditions an operator from a domain of a Banach space into itself admits a fixed point, that is

$$\underline{\Gamma}(x^*) = \underline{x}^* ,$$

this means that a method of successive substitution yields a sequence converging to a solution in the space.

This section will present the fixed point principle and some of its ramifications. A formal proof, not presented here, is available in the literature [23]-[25].

Let $\underline{\Gamma}$ be a non-linear operator whose domain and range lie in the same space. The contraction condition is stated as: an operator $\underline{\Gamma}$ is defined in $N(\underline{x}_o*, a)$ and suppose there is a number K such that for every pair of points $\underline{x}_1*$, $x_2* \in N(\underline{x}*_o, a)$ then

$$||\underline{\Gamma}(\underline{x}_2*) - \underline{\Gamma}(\underline{x}_1*)|| \leq K||\underline{x}_2* - \underline{x}_1*|| \tag{2.117}$$

where $o \leq K < 1$.

The fixed point condition is written as

$$||\underline{\Gamma}(\underline{x}_o*) - \underline{x}_o*|| < (1 - K)U \quad . \tag{2.118}$$

If the conditions of (2.117) and (2.118) are both satisfied then a unique fixed point

$$\underline{x}* \in N(\underline{x}_o*, a) \tag{2.119}$$

will exist.

A sequence can be generated via the non-linear operator $\underline{\Gamma}(\bullet)$, that is

$$\begin{aligned} \underline{\Gamma}(\underline{x}_1*) &= \underline{x}_2* \\ \underline{\Gamma}(x_2*) &= \underline{x}_3* \\ &\vdots \\ \underline{\Gamma}(\underline{x}_n*) &= \underline{x}_n*+ 1 \\ &\vdots \\ \underline{\Gamma}(\underline{x}*) &= \underline{x}* \end{aligned} \tag{2.120}$$

which it is seen terminates in a fixed point

$$\underline{\Gamma}(\underline{x}^*) = \underline{x}^*$$

The subscripts which are attached to $\underline{x}^*$ indicate the approximation to the final solution.

Then

$$\underline{x}^* = \lim_{n \to \infty} \underline{x}^*_n = \lim_{n \to \infty} \underline{\Gamma}^n(\underline{x}_o^*) ,$$

i.e. $$\left\{\underline{x}_n^*\right\} \to \underline{x}^*$$

By virtue of (2.3(d)) it is seen that (2.117) and (2.118) demand that $\underline{\Gamma}(\cdot)$ be continuous in $N(\underline{x}_o^*, a)$, and because it has a derivative of every point, in the neighborhood it is continuous. Thus (2.117) implies that the supremum of the sum of the incremental slopes will not exceed K anywhere in the neighborhood. Which in turn implies a local concept.

A method which pieces together the "local" solutions into a "global" solution and extends the region has been presented by Chu and Diaz, [14]. Trott and Christensen [11] have utilized the method in their research.

CHAPTER III

CONTINUOUS CONTROL SYSTEMS AND THE VOLTERRA SERIES

3.0 Introduction

This chapter draws upon the material derived and presented in Chapter II. Evaluating the contraction mapping and fixed point conditions are detailed in sections 3.1 and 3.2 respectively. Details of the Volterra series, its generation etc., are to be found in section 3.3.

Concluding this chapter is section 3.4, which contains a demonstration of the method which is developed in the thesis, where a differential equation representative of a practical non-linear system is solved. Previously the equation has been solved by Christensen [27] and Barrett [9], but by a different method.

3.1 Determination of the Contraction Condition

Consider the following matrix vector differential equation, which is typical of non-linear systems,

$$\dot{\underline{x}}(t) = \underline{Ax}(t) + \underline{\alpha f}(\underline{x}(t)) + \underline{By}(t). \tag{3.1}$$

The system considered is time invariant where

$\underline{A}$ = "n x n" constant matrix
$\underline{x}(t)$ = "n" vector
$\underline{y}(t)$ = "n" vector
$\underline{B}$ = "n x n" constant matrix
$\underline{\alpha}$ = "n x n" constant matrix
$\underline{f}(\underline{x}(t))$ = "n" time invariant vector

and $\underline{f}(\underline{x}(t)) \in C([a, b], [\underline{S}], R^n)$.

The solution of (3.1) is well known to be [30],

$$\underline{x}^*(t) = \underline{\Phi}(t)\underline{x}_o + \int_0^t \underline{\Phi}(t,s)\underline{By}(s)ds + \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}^*(s))ds. \tag{3.2}$$

where $\underline{\Phi}(\cdot)$ is the state transition matrix.

Similarly the solution of

$$\dot{\underline{x}}(t) = \underline{Ax}(t) + \underline{By}(t), \tag{3.3}$$

which is a time invariant linear matrix vector differential equation, is known to be, Ogata [30],

$$\underline{x}(t) = \underline{\Phi}(t)\tilde{\underline{x}}_o + \int_0^t \underline{\Phi}(t,s)\ \underline{By}(s)ds. \tag{3.4}$$

If $\underline{\alpha} = \underline{\theta}$ (the null matrix) then (3.2) takes the same form as (3.3) therefore, $\tilde{\underline{x}}_o = \underline{x}_o$.

The solution of (3.2) can be obtained by iteration techniques, which requires that a convergent sequence be developed. Let us then rewrite (3.2) and (3.4), respectively, in a simplified form, that is, let

$$\underline{x}^* = \underline{g} + \underline{\ell} + \underline{u}^* \tag{3.5}$$

$$\underline{x} = \underline{g} + \underline{\ell} \tag{3.6}$$

and substitute (3.6) into (3.5) which results in

$$\underline{x}^* = \underline{x} + \underline{u}^* , \tag{3.7}$$

which is a linear system plus a perturbation, where

$$\underline{g} = \underline{\Phi}(t)\underline{x}_o \tag{3.8}$$

$$\underline{\ell} = \int_0^t \underline{\Phi}(t,s)\ \underline{By}(s)ds \tag{3.9}$$

$$\underline{u}^* = \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}^*(s))ds. \tag{3.10}$$

If (3.10) is now substituted into (3.7) then

$$\underline{x}^*(t) = \underline{x}(t) + \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}^*(s))ds. \tag{3.11}$$

The sequence required in the sequel, and referred to above, will now be developed from (3.10) that is

$$\begin{array}{l} \underline{x}^*_1 = \underline{x} + \underline{u}_o \\ \underline{x}^*_2 = \underline{x} + \underline{u}_1 \\ \quad | \qquad\quad | \\ \quad | \qquad\quad | \\ \underline{x}^*_n = \underline{x} + \underline{u}_{n-1} \\ \quad | \qquad\quad | \\ \quad | \qquad\quad | \\ \underline{x}^* = \underline{x} + \underline{u}^* \end{array} \tag{3.12}$$

Where the subscript attached to $\underline{x}^*$ indicates the degree of approximation to the final result, that is $\underline{x}^*_1$ is the first approximation to $\underline{x}^*$.

Now substitute (3.7) into (3.10) thus

$$\underline{u}^*(t) = \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}(s) + \underline{u}^*(s))ds \tag{3.13}$$

from which it follows that

$$\underline{u}_1(t) = \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(x_1{*}(s))ds$$

$$= \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}(s) + \underline{u}_o(s))ds, \tag{3.14}$$

and if we let $\underline{u}_o(t) = \underline{\theta}$ (the null vector) then (3.14) becomes

$$\underline{u}_1(t) = \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}(s))ds, \tag{3.15}$$

where $\underline{u}_1(t)$ is the first approximation to the final solution $\underline{u}^*(t)$.

The second approximation is

$$\underline{u}_2(t) = \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}(s) + \underline{u}_1(s))ds. \tag{3.16}$$

Similarly with the third approximation, and so on up to the n^{th} approximation which is written as

$$\underline{u}_n(t) = \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}(s) + \underline{u}_{n-1}(s))ds\ . \tag{3.17}$$

If the appropriate element of (3.12) is now substituted into (3.17) then

$$\underline{u}_n(t) = \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}^*{}_n(s))ds\ . \tag{3.18}$$

It can now be seen, form (3.17), that $\underline{u}_n(t)$ is a transformation involving $\underline{u}_{n-1}(t)$, $\underline{g}(t)$ and $\underline{\ell}(t)$. Thus

$$\underline{u}_n(t) \triangleq \underline{\Gamma}_1(\underline{u}_{n-1}, \underline{g}, \underline{\ell}), \tag{3.19}$$

therefore the general term of (3.12) can be written as

$$\begin{aligned}\underline{x}^*_n &= \underline{x} + \underline{u}_{n-1} \\ &= \underline{x} + \underline{\Gamma}_1(\underline{u}_{n-2}, \underline{g}, \underline{\ell}) \\ &= \underline{\Gamma}(\underline{x}, \underline{u}_{n-2}) \\ &\triangleq \underline{\Gamma}(\underline{x}^*_{n-1})\end{aligned} \tag{3.20}$$

A new operator $\underline{\Gamma}$ has now been defined involving the approximations to the final solution and the iteration leading to the solution can be written as

$$\begin{aligned}\underline{x}^*_1 &= \underline{\Gamma}(\underline{x}_o^*) \\ \underline{x}^*_2 &= \underline{\Gamma}(\underline{x}_1^*) \\ &\;\vdots \\ \underline{x}^*_n &= \underline{\Gamma}(\underline{x}^*_{n-1}) \\ &\;\vdots \\ \underline{x}^* &= \underline{\Gamma}(\underline{x}^*) \quad .\end{aligned} \tag{3.21}$$

From the ensuing argument it is clearly seen that (3.2) is equivalent to the fixed point problem,

$$\underline{x}^* = \underline{\Gamma}(\underline{x}^*) \ , \tag{3.22}$$

and if (3.5) satisfies the contraction condition then (3.19) satisfies

$$\left\|\underline{\Gamma}(\underline{x}^*_2) - \underline{\Gamma}(\underline{x}^*_1)\right\| \leq K\left\| \underline{x}^*_2 - \underline{x}^*_1\right\| , \tag{3.23}$$

where $\underline{x}^*_1$ and $\underline{x}^*_2$ are arbitrary and $o \leq K < 1$.

From (3.20) it is seen that

$$\underline{x}^*_3 - \underline{x}^*_2 = \underline{\Gamma}(x^*_2) - \underline{\Gamma}(\underline{x}^*_1) = \underline{u}_2 - \underline{u}_1 . \tag{3.24}$$

Substitute (3.18) into (3.24) and write

$$\underline{x}^*_3 - \underline{x}^*_2 = \int_o^t \underline{\Phi}(t,s)\ \underline{\alpha}\ (\underline{f}(\underline{x}^*_2(s)) - \underline{f}(\underline{x}^*_1(s))ds. \tag{3.25}$$

But from (2.81)

$$\underline{f}(\underline{x}^*_2(s)) - \underline{f}(\underline{x}^*_1(s)) = [\frac{\partial f_j}{\partial x_k}](\underline{x}^*_2 - \underline{x}^*_1). \tag{3.26}$$

Substituting suitable elements from (3.12) into (3.26) then

$$\underline{f}(\underline{x}^*_2(s)) - \underline{f}(\underline{x}^*_1(s)) = [\frac{\partial f_j}{\partial x_k}]\ (\underline{u}_1 - \underline{u}_o) \tag{3.27}$$

and if we now substitute (3.27) into (3.25) then

$$\left|\underline{\Gamma}(\underline{x}^*_2) - \underline{\Gamma}(\underline{x}^*_1)\right| = \left|\int_o^t \underline{\Phi}(t,s)\ \underline{\alpha}\ [\frac{\partial f_j}{\partial x_k}](\underline{u}_1 - \underline{u}_o)ds\right| \tag{3.28}$$

and by considering only the R.H.S., (3.28) may be written as

$$\left|\int_o^t \underline{\Phi}(t,s)\ \underline{\alpha}\ [\frac{\partial f_j}{\partial x_k}](\underline{u}_1 - \underline{u}_o)ds\right| \leq \int_o^t\left|\underline{\Phi}(t,s)\right|ds$$

$$\sup_t\left|\ \underline{\alpha}\ [\frac{\partial f_j}{\partial x_k}]\right|\left\|\underline{u}_1 - \underline{u}_o\right\| . \tag{3.29}$$

To make that which is to follow, notationally simple it will be necessary to make certain substitutions. Initially let

$$\underline{P}(t) = \int_0^t \left|\underline{\Phi}(t,s)\right| ds \ , \tag{3.30}$$

where $\underline{P}(\cdot)$ is an "n x n" matrix and

$$\int_0^t \left| \quad \right| ds > o \ \forall \ t > o \tag{3.31}$$

by hypothesis.

Then by definition

$$P_{ik}(t) = \int_0^t \left|\Phi_{ik}(t,s)\right| ds \tag{3.32}$$

therefore,

$$P_{ik}(t_1) = \int_0^{t_1} \left| \quad \right| ds \tag{3.33}$$

and

$$P_{ik}(t_2) = \int_0^{t_2} \left| \quad \right| ds \tag{3.34}$$

and if $t_2 > t_1$ then

$$P_{ik}(t_2) = \int_0^{t_1} \left| \quad \right| ds + \int_{t_1}^{t_2} \left| \quad \right| ds \ . \tag{3.35}$$

Now substitute (3.33) and (3.34) into (3.35) and write

$$P_{ik}(t_2) = P_{ik}(t_1) + \int_{t_1}^{t_2} \left| \quad \right| ds \ . \tag{3.36}$$

The original hypothesis (3.31) means that

$$\int_{t_1}^{t_2} \left| \quad \right| ds > o \tag{3.37}$$

providing $t_2 > t_1$,

thus (3.36) can be rewritten as

$$P_{ik}(t_2) \geq P_{ik}(t_1) \tag{3.38}$$

and by induction

$$P_{ik}(t_n) \geq P_{ik}(t_{n-1}) \quad (t_n > t_{n-1}). \tag{3.39}$$

Therefore, it can be concluded that

$$\underset{t=\infty}{P_{ik}(t)} = P_{ik} \geq P_{ik}(t) \quad \forall t<\infty \tag{3.40}$$

where P_{ik} is the supremum of the terms which satisfy (3.40).

The $\underline{P}$ matrix is therefore dominated by a matrix of real numbers, each of which is independent of t. Thus

$$\begin{bmatrix} P_{11}(t) & P_{12}(t) \\ P_{21}(t) & P_{22}(t) \end{bmatrix} \leq \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \tag{3.41}$$

for a 2^{nd} order system, and (3.30) can be written as

$$\underline{P} = \int_0^{t=\infty} \left| \underline{\Phi}(t,s) \right| ds \tag{3.42}$$

where

$$P_{ik} = \int_0^{t=\infty} \left| \Phi_{ik}(t,s) \right| ds \tag{3.43}$$

and t is the time interval under consideration that is, $o \leq t \leq \infty$ and s is the time elapsed, therefore $o \leq s \leq t$.

Return again to (3.29) and let

$$\underline{z}(t) = \sup_t \left| \underline{\alpha}[\frac{\partial fj}{\partial x_k}] \right| \tag{3.44}$$

where $\underline{z}(\cdot)$ is a "n x n" matrix, and a component of this matrix is written as

$$z_{km}(t) = \sup_t \left| \alpha_{kj} \frac{\partial fj}{\partial x_k} \right| , \tag{3.45}$$

where $1 \leq m \leq j$.

Then (3.29) can be rewritten as

$$\left| \int_0^t \underline{\Phi}(t,s) \; \underline{\alpha} \; [\frac{\partial f_j}{\partial x_k}] (\underline{u}_1 - \underline{u}_o) ds \right| \leq \sup_i \left| \sum_{k=1}^{n} P_{ik} \; z_{km}(t) \right| \| \underline{u}_1 - \underline{u}_o \| \tag{3.46}$$

and by the original definition of the norm (see equations 2.6 and 2.7)

$$\sup_i \left| \sum_{k=1}^{n} P_{ik} \; z_{km}(t) \right| = \| \underline{P} \; \underline{z}(t) \| . \tag{3.47}$$

Hence by substituting (3.42) into (3.41) it is found that

$$\left| \int_0^t \underline{\Phi}(t,s) \; \underline{\alpha} \; [\frac{\partial f_j}{\partial x_k}] (\underline{u}_1 - \underline{u}_o) ds \right| \leq \| \underline{P} \; \underline{z}(t) \| \; \| \underline{u}_1 - \underline{u}_o \| , \tag{3.48}$$

now substitute (3.48) into (3.28) to obtain

$$\left| \underline{\Gamma}(\underline{x}_2 *) - \underline{\Gamma}(\underline{x}_1 *) \right| \leq \| \underline{P} \; \underline{z}(t) \| \; \| \underline{u}_1 - \underline{u}_o \| . \tag{3.49}$$

Substituting elements of (3.12) into (3.49) causes

$$\|\underline{u}_1 - \underline{u}_o\| = \|\underline{x}_2{}^* - \underline{x}_1{}^*\| \tag{3.50}$$

therefore (3.49) can be rewritten as

$$\left|\underline{\Gamma}(\underline{x}_2{}^*) - \underline{\Gamma}(\underline{x}_1{}^*)\right| \leq \|\underline{P}\ \underline{z}(t)\|\ \|\underline{x}_2{}^* - \underline{x}_1{}^*\| \ . \tag{3.51}$$

Taking the supremum of (3.51) over t results in

$$\sup_t \left|\underline{\Gamma}(\underline{x}_2{}^*) - \underline{\Gamma}(\underline{x}_1{}^*)\right| \leq \sup_t \|\underline{P}\ \underline{z}(t)\|\ \|\underline{x}_2{}^* - \underline{x}_1{}^*\| \tag{3.52}$$

and since

$$\sup |\cdot| = \|\cdot\| \tag{3.53}$$

then clearly (3.52) can be written as

$$\|\underline{\Gamma}(\underline{x}_2{}^*) - \underline{\Gamma}(\underline{x}_1{}^*)\| \leq \sup_t \|\underline{P}\ \underline{z}(t)\|\ \|\underline{x}_2{}^* - \underline{x}_1{}^*\| \ , \tag{3.54}$$

which is of the same form as (3.23) if

$$\sup_t \|\underline{P}\ \underline{z}(t)\| = K < 1; \tag{3.55}$$

which satisfies the contraction condition.

A natural question to raise at this point is, how does the F derivative fit into the picture? The connection between the F

derivative and the contraction condition will now be shown.

Let $\underline{x}_1$, $\underline{x}_2$ ε $\underline{S} \subset \underline{X}$, where $\underline{S}$ is a subspace of the Banach space $\underline{X}$ and $\underline{x}_1$, $\underline{x}_2$ are arbitrary in $\underline{S}$, also let $\underline{G}$: $\underline{S} \subset \underline{X} \to \underline{X}$. The references for this section are [18], [19] and [26] who have utilized the mapping function $\underline{G}$ as follows,

$$\|\underline{G}(\underline{x}_2) - \underline{G}(\underline{x}_1)\| \le \sup_{o<\gamma<1} \left\|\underline{G}'(\underline{x}_1 + \gamma(\underline{x}_2 - \underline{x}_1))\right\| \, \|\underline{x}_2 - \underline{x}_1\|. \qquad (3.56)$$

Let us now apply this inequality to the mapping

$$\underline{G}\,\underline{w} = \underline{\Gamma}\,\underline{w} - \underline{\Gamma}'(\underline{x})\underline{w} \qquad (3.57)$$

where $\underline{x}$ is fixed in $\underline{S}$ and $\underline{w} \varepsilon \underline{S}$.

Differentiating (3.57) with respect to $\underline{w}$ gives

$$\underline{G}'(\underline{w}) = \underline{\Gamma}'(\underline{w}) - \underline{\Gamma}'(\underline{x}) \qquad (3.58)$$

now replace $\underline{w}$ in (3.57) by $\underline{x}_1$ and then by $\underline{x}_2$ and subtract one equation from the other, which results in

$$\underline{G}(\underline{x}_2) - \underline{G}(\underline{x}_1) = \underline{\Gamma}(\underline{x}_2) - \underline{\Gamma}(\underline{x}_1) - \underline{\Gamma}'(\underline{x})(\underline{x}_2 - \underline{x}_1) \qquad (3.59)$$

and taking the norm of (3.59) gives

$$\|\underline{G}(\underline{x}_2) - \underline{G}(\underline{x}_1)\| = \|\underline{\Gamma}(\underline{x}_2) - \underline{\Gamma}(\underline{x}_1) - \underline{\Gamma}'(\underline{x})(\underline{x}_2 - \underline{x}_1)\|. \qquad (3.60)$$

Now replace w in (3.58) by $(\underline{x}_1 + \gamma(\underline{x}_2 - \underline{x}_1))$ and take the norm of the result

$$\|\underline{G}'(\underline{x}_1 + \gamma(\underline{x}_2 - \underline{x}_1)\| = \|\underline{\Gamma}'(x_1 + \gamma(\underline{x}_2 - \underline{x}_1)) - \underline{\Gamma}'(\underline{x})\| , \tag{3.61}$$

and substitute (3.60) and (3.61) into (3.56)

$$\|\underline{\Gamma}(\underline{x}_2) - \underline{\Gamma}(\underline{x}_1) - \underline{\Gamma}'(\underline{x})(\underline{x}_2 - \underline{x}_1)\| \leq \sup_{o<\gamma<1} \|\underline{\Gamma}'(\underline{x}_1 + \gamma(\underline{x}_2 - \underline{x}_1)) - \underline{\Gamma}'(\underline{x})\| \, \|\underline{x}_2 - \underline{x}_1\| \tag{3.62}$$

which can also be written as

$$\|\underline{\Gamma}(\underline{x}_2) - \underline{\Gamma}(\underline{x}_1)\| \leq \sup_{o<\gamma<1} \|\underline{\Gamma}'(\underline{x}_1 + \gamma(\underline{x}_2 - \underline{x}_1))\| \, \|\underline{x}_2 - \underline{x}_1\| . \tag{3.63}$$

Reference [18] page 660, has shown that

$$\sup_{o<\gamma<1} \|\underline{\Gamma}'(\underline{x}_1 + \gamma(\underline{x}_2 - \underline{x}_1))\| \leq \sup_{\underline{x} \, \varepsilon \, \underline{s}} \|\underline{\Gamma}'(\underline{x})\| \tag{3.64}$$

and if

$$\sup_{\underline{x} \, \varepsilon \, \underline{s}} \|\underline{\Gamma}'(\underline{x})\| = K < 1 \tag{3.65}$$

then (3.63) can be written as

$$\|\underline{\Gamma}(\underline{x}_2) - \underline{\Gamma}(\underline{x}_1)\| \leq \sup_{o<\gamma<1} \|\underline{\Gamma}'(\underline{x}_1 + \gamma(\underline{x}_2 - \underline{x}_1))\| \, \|\underline{x}_2 - \underline{x}_1\|$$

$$\leq \sup_{\underline{x}\varepsilon\underline{S}} \|\underline{\Gamma}'(\underline{x})\| \, \|\underline{x}_2 - \underline{x}_1\|$$

$$\leq K\|\underline{x}_2 - \underline{x}_1\| . \tag{3.66}$$

Thus an upper bound is provided for $\underline{\Gamma}(\underline{x}_2*) - \underline{\Gamma}(\underline{x}_1*)$ in terms of $\underline{\Gamma}'$. Therefore, it can be concluded that if the norm of the F derivative of the operator $\underline{\Gamma}$ is less than one, then the contraction condition is satisfied.

It has been shown in (3.66) that the contraction condition is satisfied if the norm of the F derivative is less than one. But how is this norm to be evaluated in practice? It will now be shown that an equality does exist between (3.55) and (3.65), to some degree, and that it is possible to obtain a practical estimate of the norm of the F derivative by determining the norm of

$$\|\underline{P}\ \underline{z}(t)\| \tag{3.67}$$

and limiting its magnitude to be less than one.

The space being utilized here is the function space $C([a,b],\underline{S},R^n)$, which was described in section 2.1 and specifies the domain of operation. Equation (3.2) is rewritten for convenience

$$\underline{x}*(t) = \underline{\Phi}(t)\underline{x}_o + \int_0^t \underline{\Phi}(t,s)\ \underline{B}\ \underline{y}(s)ds + \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}\ (\underline{x}*(s))ds, \tag{3.68}$$

which is equivalent to the fixed point problem

$$\underline{x}^* = \underline{\Gamma}(\underline{x}^*) \tag{3.69}$$

on C([a,b], $\underline{S}$). Thus $\underline{\Gamma}(\cdot)$ is a mapping where

$$\underline{\Gamma}(\underline{x}^*) = \underline{\Phi}(t)\ \underline{x}_o + \int_0^t \underline{\Phi}(t,s)\ \underline{B}\ \underline{y}(s)ds + \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}^*(s))ds \tag{3.70}$$

as previously stated.

To determine the F derivative of $\underline{\Gamma}$ it is necessary to utilize Lemma 8.5 of Falb and DeJong [15], which is restated here for convenience.

(i) Let $\underline{S}$ be an open domain in R^n, and T be an open set in R^1, which contains [a,b]. Suppose that $\underline{T}(t,\underline{x},s)$ is a map of $T \times \underline{S} \times T$ into R^n and is measurable in s for each fixed $\underline{x}(\cdot)$ and t, and is continuous in $\underline{x}(\cdot)$ for each fixed s and t.

Where:

t is the time interval under consideration.

s is the time elapsed and $o \underline{<} s \underline{<} t$.

What is implied by $\underline{T}(\cdot)$ being measurable in s for each fixed t? If $\underline{x}(\cdot)$ is a type of continuous function then t is the time over which the function is being considered, the time of duration. If the elapsed time and final time are fixed then $\underline{x}(\cdot)$ is continuous by hypothesis.

(ii) $\dfrac{\partial \underline{T}(\cdot)}{\partial \underline{x}}$ is a map of T x $\underline{S}$ x T into $L(R^n, R^n)$ (for details of this space see [15], [16] and [18] and definition 19). The partial derivative is measurable in s for each fixed $\underline{x}(\cdot)$ and t, and continuous in $\underline{x}(\cdot)$ for each fixed s and t.

(iii) There is an integrable function m(t,s) of s with

$$\sup_t \int_a^b m(t,s)\,ds < \infty \tag{3.71}$$

such that $\| \underline{T}(t,\underline{x},s) \| \leq m(t,s)$ and $\left\| \dfrac{\partial \underline{T}(t,\underline{x},s)}{\partial \underline{x}} \right\| \leq m(t,s)$ on $T \times \underline{S} \times T$.

(3.72)

(iv) $$\lim_{t \to t'} \int_a^b \| \underline{T}(t,\underline{x},s) - \underline{T}(t',\underline{x},s) \|\, ds = o \tag{3.73}$$

and

$$\lim_{t \to t'} \int_a^b \left\| \frac{\partial \underline{T}(t,\underline{x}(s),s)}{\partial \underline{x}} - \frac{\partial \underline{T}(t',\underline{x}(s),s)}{\partial \underline{x}} \right\| ds = o \tag{3.74}$$

for all t, t' in [a,b] and $\underline{x}(\cdot)$ in $C([a,b],\underline{S})$. Thus the mapping $\underline{T}$ given by

$$\underline{\Gamma}(u(t)) = \int_a^b \underline{T}(t, u(s), s)\,ds \tag{3.75}$$

is a differentiable mapping of $C([a,b], \underline{S})$ into $C([a,b], R^n)$ with the derivative of the mapping being given by

$$(\underline{\Gamma}'\underline{u}\ \underline{V})(t) = \int_a^b \frac{\partial \underline{T}(t, \underline{u}(s), s)}{\partial \underline{u}}\ \underline{V}(s)\,ds \tag{3.76}$$

where $\underline{u}(\cdot) \in C([a,b], \underline{S})$.

The proof, which is contained in [15] is not reproduced here, but the argument presented by the authors proceeds as follows

(i) observes that $\underline{\Gamma}$ and $\underline{\Gamma}'$ are actually mapping functions.

(ii) establishes the lemma by utilizing these mapping functions, the mean value theorem and the Lebesque dominated-convergence theorem.

Equations (2.20), (2.21) and (2.22) of the thesis are similar to those used and proved by Falb and DeJong in the establishment of this lemma.

Therefore, if

$$\underline{T}(t, \underline{x}(s), s) = \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}(s)) \tag{3.77}$$

and the limits of the integration are set, that is a = o and b = t then

$$\underline{\Gamma}'(\underline{x}(t)) = \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha} \left[\frac{\partial f_j}{\partial x_k}\right] \underline{x}(s)ds \tag{3.78}$$

where

$\left[\frac{\partial f_j}{\partial x_k}\right]$ is a Jacobian matrix

and $\underline{\Gamma}'(\underline{x}(t))$ is the Frechet derivative.

The norm of the F derivative can be written as

$$||\underline{\Gamma}'|| = \sup_{||\underline{x}|| \leq 1} ||\underline{\Gamma}'(\underline{x}(t))|| \tag{3.79}$$

$$= \sup_{||\underline{x}|| \leq 1}\ \sup_i\ \sup_t \left|\sum_{j=1}^{n} \sum_{k=1}^{n} \int_0^t \underline{\Phi}_{ik}(t,s)[\alpha_{kj}][\frac{\partial f_j}{\partial x_i}]x_i(s)ds\right|. \tag{3.80}$$

This is often difficult to evaluate. Therefore, a coarser estimate of the norm is taken in practice, and that is

$$\leq \sup_i \left(\sum_{k=1}^{n} \int_o^t \left| \Phi_{ik}(t,s) \right| ds \ \sup_t \sum_{j=1}^{n} \left| \alpha_{kj} \frac{\partial f_j}{\partial x_i} \right| \| \underline{x} \| \right) \tag{3.81}$$

which can be called

$$\| \underline{\Gamma}'(\underline{x}) \|_C \tag{3.82}$$

where the subscript "C" means coarse to some degree. The above is for the continuous system.

It is well known that the derivative is a result of a limiting process. It is easily shown from

$$\left\| \frac{\underline{\Gamma}(\underline{x}_2^*) - \underline{\Gamma}(\underline{x}_1^*)}{x_2^* - \underline{x}_1^*} - \underline{\Gamma}'(\underline{x}_1^*) + \underline{\Gamma}'(\underline{x}_1^*) \right\|$$

$$= \left\| \frac{\underline{\Gamma}(\underline{x}_2^*) - \underline{\Gamma}(\underline{x}_1^*)}{\underline{x}_2^* - \underline{x}_1^*} \right\| \tag{3.83}$$

that,

$$\left\| \frac{\underline{\Gamma}(\underline{x}_2^*) - \underline{\Gamma}(\underline{x}_1^*)}{\underline{x}_2^* - \underline{x}_1^*} \right\| \leq \left\| \frac{\underline{\Gamma}(\underline{x}_2^*) - \underline{\Gamma}(\underline{x}_1^*)}{\underline{x}_2^* - \underline{x}_1^*} - \underline{\Gamma}'(\underline{x}_1) \right\| + \| \underline{\Gamma}'(\underline{x}_1) \| \tag{3.84}$$

In the definition of the F derivative, Chapter 2, it was shown that

$$\left\| \frac{\underline{\Gamma}(\underline{x}_2^*) - \underline{\Gamma}(\underline{x}_1^*)}{\underline{x}_2^* - \underline{x}_1^*} - \underline{\Gamma}'(\underline{x}_1^*) \right\| = o \ , \tag{3.85}$$

in the limit. Thus,

$$\lim_{\substack{\underline{x}_2^* \to \underline{x}_1^* \\ \underline{x}_1^*,\ \underline{x}_2^* \in \underline{s}}} \left\| \frac{\underline{\Gamma}(\underline{x}_2^*) - \underline{\Gamma}(\underline{x}_1^*)}{\underline{x}_2^* - \underline{x}_1^*} \right\| = \left\| \underline{\Gamma}'(\underline{x}_1^*) \right\| \tag{3.86}$$

and

$$\sup_{\underline{x}_1^*,\ \underline{x}_2^* \in \underline{s}} \left\| \frac{\underline{\Gamma}(\underline{x}_2^*) - \underline{\Gamma}(\underline{x}_1^*)}{\underline{x}_2^* - \underline{x}_1^*} \right\| \leq \sup_{\underline{x}^* \in \underline{S}} \left\| \underline{\Gamma}'(\underline{x}^*) \right\| . \tag{3.87}$$

alternatively

$$\sup_{\substack{o<\gamma<1 \\ \underline{x}_1^*,\ \underline{x}_2^* \in \underline{S}}} \left\| \underline{\Gamma}'(\underline{x}_1^* + \gamma(\underline{x}_2^* - \underline{x}_1^*)) \right\| \leq \sup_{\underline{x}^* \in \underline{S}} \left\| \underline{\Gamma}'(\underline{x}^*) \right\| \tag{3.88}$$

It is clearly seen that the L.H.S. of the inequalities (3.87) and (3.88) are estimates of the norm of the actual derivative. The magnitude of the norm of the F derivative is required to prove that a contraction condition does exist. This was the conclusion that was reached following (3.66). The practical determination of the fact, that a contraction condition does actually exist, is vital if a Volterra series solution is to be found for the problem. Details of the Volterra series are to be found in section 3.3. Thus a practical method of evaluating the norm of the F derivative is required, this can be achieved as follows. It can be shown from (3.87) and (3.82) that

$$\begin{aligned} \left\| \frac{\underline{\Gamma}(\underline{x}_2^*) - \underline{\Gamma}(\underline{x}_1^*)}{\underline{x}_2^* - \underline{x}_1^*} \right\| &\leq \sup_{\underline{x}^* \in \underline{S}} \left\| \underline{\Gamma}'(\underline{x}^*) \right\| \\ &\leq \sup \left\| \underline{\Gamma}'(\underline{x}^*) \right\|_c \end{aligned} \tag{3.89}$$

and in (3.54) it was shown that

$$\left\| \frac{\underline{\Gamma}'(\underline{x}_2^*) - \underline{\Gamma}(\underline{x}_1^*)}{\underline{x}_2^* - \underline{x}_1^*} \right\| \leq \sup_t \left\| \underline{P}\ \underline{z}(t) \right\| . \tag{3.90}$$

Therefore, if the arguments which terminated in (3.54) and (3.81) are compared it is clearly seen that (3.55) can be equated to (3.82), that is,

$$\begin{aligned} \left\| \underline{\Gamma}'(\underline{x}^*) \right\|_c &= \sup_t \left\| \underline{P}\ \underline{z}(t) \right\| = K < 1 \\ &= \sup_i \ \sup_t \left| \sum_{k=1}^{n} P_{ik}\ z_{km}(t) \right| = K < 1 . \end{aligned} \tag{3.91}$$

It is conceded that this is a coarse estimate of the norm of the F derivative, but it is easily evaluated for practical systems. Therefore, it can be concluded that if the magnitude of the coarse norm is limited to be smaller than one then the L.H.S. of (3.89) is also less than one. The magnitude of the norm is controlled by limiting the size of $\underline{S}$, more details of this are to be found in section 3.3.

3.2 Determination of the Fixed Point Condition

From (3.19) and (2.118) the fixed point condition is written as

$$\left\| \underline{\Gamma}_1(\underline{u}_o, \underline{g}, \underline{\ell}) - \underline{u}_o \right\| < (1 - K)\ U \tag{3.92}$$

which can be rewritten as

$$\|\underline{u}_1 - \underline{u}_o\| < (1 - K)U, \tag{3.93}$$

and by the definition of norm

$$\|\underline{u}_1 - \underline{u}_o\| = \sup_t \|\underline{u}_1(t) - \underline{u}_o(t)\| < (1 - K)U. \tag{3.94}$$

If $\underline{u}_o(t) = \underline{\theta}$ and (3.14) is substituted into (3.94) then

$$\sup_t \left| \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}^*(s))ds \right| < (1 - K)U \tag{3.95}$$

from which it follows that

$$\begin{aligned} \left| \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}^*(s))ds \right| &\leq \int_0^t |\underline{\Phi}(t,s)|\,\left|\ \underline{\alpha}\ \underline{f}(\underline{x}^*(s))\right| ds \\ &\leq \int_0^t |\underline{\Phi}(t,s)| ds \ \sup_t \left| \underline{\alpha}\ \underline{f}(\underline{x}^*(t)) \right|, \end{aligned} \tag{3.96}$$

now utilizing (3.43) again, that is let

$$P_{ik} = \int_0^{t=\infty} |\Phi_{ik}(t,s)| ds \tag{3.97}$$

and

$$L_k = \sup_t \left| \alpha_{kj} f_j(\underline{x}^*(t)) \right|. \tag{3.98}$$

Thus (3.95) may be written as

$$\left|\int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}^*(s))ds\right| \leq \sup_i \left|P_{ik}L_k\right| , \tag{3.99}$$

and by the definition of norm

$$\sup_i \left|P_{ik}L_k\right| = \left\|\underline{P}\ \underline{L}(t)\right\| \tag{3.100}$$

therefore,

$$\sup_t \left|\int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}^*(s))ds\right|$$

$$\leq \sup_t \left\|\underline{P}\ \underline{L}(t)\right\| . \tag{3.101}$$

Hence by comparing (3.101) and (3.95) it is easily seen that if

$$\sup_t \left\|\underline{P}\ \underline{L}(t)\right\| < (1 - K)U \tag{3.102}$$

then the fixed point condition is satisfied.

3.3 The Volterra Series

A Volterra series is a functional polynomial, that is a type of power series which gives the output explicitly in terms of the input. The series is based upon the Weierstrass theorem which states "any function continuous in a closed interval can be uniformly approximated within any prescribed tolerance, over the interval, by some polynomial - erected from linear elements".

The use of the Volterra series as a representation of non-linear systems is justified by the fact that if the output of a system is continuously dependent upon its input, and if the input takes its values from a bounded subset of a normed linear space, then the output can be uniformly approximated by a Volterra series. If the series converges then the system exhibits B.I.B.O. stability. The deduction that a system is B.I.B.O. stable via convergence is important. A convergence criterion has been developed by Christensen, [10], based on the contraction mapping principle and the definition of an analytic system. This criterion will be suitably modified to suit the vector matrix form utilized here, and will be presented later in this section.

It has also been shown, by Trott and Christensen [33], that the Volterra series is unique provided certain conditions are satisfied:

(1) The norm of the input is bounded

(2) The fixed point condition holds

(3) The contraction condition holds.

A natural question is can the solution of (3.1) be expressed as a Volterra series? To show that the answer to the above question is in the affirmative, for a class of systems, a Volterra series will now be generated via (3.19).

By combining (3.54) and (3.55) and then substituting elements of (3.12) into the result gives

$$\| \underline{\Gamma}_1(\underline{u}_2) - \underline{\Gamma}_1(\underline{u}_1) \| \leq K \| \underline{u}_2 - \underline{u}_1 \| , \tag{3.103}$$

and from (3.19)

$$\underline{u}_3 = \underline{\Gamma}_1(\underline{u}_2)$$

$$\underline{u}_2 = \underline{\Gamma}_1(\underline{u}_1)$$

and by substituting the above equations into (3.103) results in

$$\|\underline{u}_3 - \underline{u}_2\| \leq K\|\underline{u}_2 - \underline{u}_1\| , \tag{3.104}$$

and by similar reasoning

$$\|\underline{u}_2 - \underline{u}_1\| \leq K\|\underline{u}_1 - \underline{u}_o\| . \tag{3.105}$$

By applying the axiom, $\|a - b\| \geq \|a\| - \|b\|$ to (3.105) gives

$$\|\underline{u}_2\| \leq K\|\underline{u}_1 - \underline{u}_o\| + \|\underline{u}_1\| ; \tag{3.106}$$

similarly

$$\|\underline{u}_3\| \leq K\|\underline{u}_2 - \underline{u}_1\| + \|\underline{u}_2\|. \tag{3.107}$$

Now substitute (3.105) and (3.106) into (3.107) then,

$$\|\underline{u}_3\| \leq K^2\|\underline{u}_1 - \underline{u}_o\| + K\|\underline{u}_1 - \underline{u}_o\| + \|\underline{u}_1\| \tag{3.108}$$

A general form of equation (3.108) may now be written as

$$\|\underline{u}_n\| \leq \left(\sum_{i=1}^{n-1} K^i\|\underline{u}_1 - \underline{u}_o\|\right) + \|\underline{u}_1\| . \tag{3.109}$$

By adding $\|\underline{u}_o\|$ to each side of (3.109) and by applying the axiom $\|a - b\| \leq \|a\| + \|b\|$; then (3.109) may be rewritten as

$$\|\underline{u}_n - \underline{u}_o\| \leq \left(\sum_{i=1}^{n-1} K^i \| \underline{u}_1 - \underline{u}_o \| \right) + \|\underline{u}_1 - \underline{u}_o\|$$

$$\leq \| \underline{u}_1 - \underline{u}_o \| (1 + k + K^2 + \text{----} + K^{n-1}) < U \qquad (3.110)$$

where

$$U = \|\underline{u}\| \leq \sup_i \ \sup_t \left| u_i(t) \right|, \qquad (3.111)$$

which is the closure of the region in which the solution $\underline{u}^*$ is to be found.

Since $(1 + K + K^2 + \text{----} + K^{n-1})$ is a geometric series the sum over "n" terms is

$$S_n = \sum_{i=0}^{n-1} K^i = \frac{1 - K^n}{1 - K} \qquad (3.112)$$

and in the limit as "n" $\to \infty$

$$\lim_{n\to\infty} \frac{1 - K^n}{1 - K} = \frac{1}{1-K}, \qquad (3.113)$$

because $o<K<1$

now substitute (3.113) into (3.110) and let $\underline{u}^* = \lim_{n\to\infty} \underline{u}_n$ then

$$\| \underline{u}^* - \underline{u}_o \| \leq \frac{1}{1-K} \|\underline{u}_1 - \underline{u}_o\| < U \qquad (3.114)$$

which implies that

$$||\underline{u}_1 - \underline{u}_o|| < (1 - K)\ U. \tag{3.115}$$

If (3.115) is compared with (2.118) it is easily seen that the inequality (3.115) is a statement of the fixed point condition.

It will now be shown that the power series (3.110) is convergent. Rewrite (3.110) in series form as follows

$$||\underline{u}_n - \underline{u}_o|| \leq ||\underline{u}_1 - \underline{u}_o|| + K\,||\underline{u}_1 - \underline{u}_o|| + \text{------}$$
$$\text{----} + K^{n-1}||\underline{u}_1 - \underline{u}_o|| \tag{3.116}$$

now if (3.104) and (3.105), and so on, are substituted into (3.116) it can be rewritten as

$$||\underline{u}_n - \underline{u}_o|| \leq ||\underline{u}_1 - \underline{u}_o|| + ||\underline{u}_2 - \underline{u}_1|| + \text{----}$$
$$\text{---} + ||\underline{u}_n - \underline{u}_{n-1}|| \tag{3.117}$$

and if the axiom $||a-b|| \geq ||a|| - ||b||$ is applied to the L.H.S. of (3.117) then,

$$||\underline{u}_n|| \leq ||\underline{u}_o|| + ||\underline{u}_1 - \underline{u}_o|| + ||\underline{u}_2 - \underline{u}_1|| + \text{---}$$
$$\text{----} + ||\underline{u}_n - \underline{u}_{n-1}||\,, \tag{3.118}$$

therefore,

$$\underline{u}_n = \underline{u}_o + (\underline{u}_1 - \underline{u}_o) + (\underline{u}_2 - \underline{u}_1) + ----$$

$$---- + (u_n - u_{n-1}) \tag{3.119}$$

and in the limit as "n"→∞

$$\underline{u}^* = \underline{u}_o + (\underline{u}_1 - \underline{u}_o) + (\underline{u}_2 - \underline{u}_1) + ----$$

$$---- + (\underline{u}_n - \underline{u}_{n-1}) + ---- ; \tag{3.120}$$

which is the form of a convergent series.

It will now be shown that the power series (3.110) converges in a region where o<K<1. For details of power series and associated topics see typically Courant and John [28], Lick [29].

Substitute (3.105) into (3.104) and write

$$||\underline{u}_3 - \underline{u}_2|| \leq K^2 || \underline{u}_1 - \underline{u}_o|| \tag{3.121}$$

then by induction

$$||\underline{u}_n - \underline{u}_{n-1}|| \leq K^{n-1}||\underline{u}_1 - \underline{u}_o|| , \tag{3.122}$$

and by substituting (3.115) into (3.122) results in

$$||\underline{u}_n - \underline{u}_{n-1}|| < K^{n-1}(1 - K)\ U. \tag{3.123}$$

Now apply the ratio test to (3.118) via inequality (3.123) to prove that the series (3.119) does indeed converge, that is,

$$\frac{K^n(1 - K)U}{K^{n-1}(1 - K)U} = K \ , \tag{3.124}$$

and because $K<1$ the series converges.

Thus equation (3.19) has a unique series solution in the region where the following two conditions hold.

(1) From (3.103) i.e.

$$\|\underline{\Gamma}_1(\underline{u}_2) - \underline{\Gamma}_1(\underline{u}_1)\| \leq K\|\underline{u}_2 - \underline{u}_1\| \tag{3.125}$$

and

(2) From (3.115) i.e. $\|\underline{u}_1 - \underline{u}_o\| < (1 - K)U$, and because of equation (3.19) this can be rewritten as

$$\|\underline{\Gamma}_1(\underline{u}_1) - \underline{u}_o\| < (1 - K)U. \tag{3.126}$$

clearly

$$(1 - K)U \tag{3.127}$$

is the region of uniqueness of the solution. Thus the series (3.120) has been shown to be convergent in the region where $o<K<1$.

One final point remains to be developed, before it can be shown that equation (3.1) can be cast into the form of a Volterra series, and that is a "convergence criterion".

Therefore, we quite naturally ask the following question, "When can it be inferred that a Volterra series exists and is convergent for a given system if the contraction mapping principle shows that system to exhibit B.I.B.O. stability?"

Christensen [10] has answered this question for the single input-single output system, his method will be extended here to cover multidimensional systems.

The answer to the question posed above is simply, that the given system in addition to having a unique solution must also be analytic [6], [7], [35]-[37].

Specifically, in vector form the Volterra series is written as

$$\underline{x}(t) = \underline{g} + \int_{-\infty}^{\infty}\underline{h}(\tau)\underline{y}(t-\tau)d\tau +$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\underline{h}_2(\tau_1,\tau_2)\underline{y}(t-\tau_1)\underline{y}(t-\tau_2)d\tau_1 d\tau_2 + ---$$

$$--- + \int_{-\infty}^{\infty}--\int_{-\infty}^{\infty}\underline{h}_n(\tau_1,----\tau_n)\underline{y}(t-\tau_1) -----$$

$$-- \underline{y}(t-\tau_n)d\tau_1 ---- d\tau_n \tag{3.128}$$

where

$\underline{g}$ = initial conditions on $\underline{x}(t)$ and its derivatives

$\underline{y}(t)$ = system input

$\underline{x}(t)$ = system output

which gives the output $\underline{x}(t)$ for a given input $\underline{y}(t)$ for the system in question.

Brilliant [7] has shown that if

$$\|\underline{h}_i\| = \int_{-\infty}^{\infty} \text{--} \int_{-\infty}^{\infty} \left|\underline{h}_i(\tau_1, \text{----}, \tau_i)\right| d\tau_1 \text{ --- } d\tau_i < \infty$$

for all i, such that (3.230) can be cast into the form

$$\underline{\overline{X}}^* \leq \sum_{n=o}^{\infty} \|\underline{h}_n\| \, \underline{Y}^n \qquad (3.129)$$

and if the right hand side of (3.129) is a convergent series, then the given system is analytic.

Now consider the series generated in (3.12). This series refers to system (3.1) and it is readily seen that, with the aid of the material developed in this section, the series can be cast into the form of (3.129), and is convergent in the region considered. Thus the system which is represented by equation (3.1) is an analytic system.

In conclusion, consider once again equation (3.1), if a solution is obtained for the equation by the use of the contraction mapping principle, then this solution is unique. If, in addition, this solution can be cast into the form of (3.129) then, (3.1) represents an analytic system. Thus, it can be stated that a Volterra series solution exists for the system.

It will now be shown that equation (3.1) can be cast into the form of a Volterra series. Substitute (3.15) in (3.16) and write

$$\underline{u}_2(t) = \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}(s_1) + \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}(s_2))ds_1 ds_2 \quad (3.130)$$

sinilarly with $\underline{u}_3(t)$, etc.

Now substitute (3.120) into equation (3.5) and write

$$\underline{x}^*(t) = \underline{g} + \ell + \underline{u}_o + (\underline{u}_1 + \underline{u}_o) + \text{-----}$$

$$\text{---} + (\underline{u}_n - \underline{u}_{n-1}) + \text{-----}$$

$$= \underline{\Phi}(t)\underline{x}_o + \int_0^t \underline{\Phi}(t,s)\ \underline{B}\ \underline{y}(s)ds \quad (3.131)$$

$$+ \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(x(s))ds + \text{-----}$$

$$\text{---} + \text{------},$$

where

$$\underline{g} = \underline{\Phi}(t)\underline{x}_o \quad (3.8)$$

$$\underline{\ell} = \int_0^t \underline{\Phi}(t,s)\ \underline{B}\ \underline{y}(s)ds \quad (3.9)$$

$$\underline{u}_o = \theta \text{ (null vector)}$$

$$\underline{u}_1 - \underline{u}_o = \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}(s))ds.$$

Similarly with $\underline{u}_2 - \underline{u}_1$ and so on.

Clearly equation (3.131) is the Volterra series.

Conclusion:

This section has clearly shown that for a stable non-linear system we can represent the output as a convergent power series if the contraction mapping and fixed point conditions are satisfied. Furthermore, if the system equation is analytic the resulting power series representing the system output is a Volterra series.

3.4 The Solution of a Non-linear System

To illustrate and amplify the method which has been developed in the preceeding pages, the solution to a typical problem will now be offered. Previously, this same problem has been solved by Christensen [26], but by a different method. It will be interesting, in the sequel, to compare the results previously obtained, with those obtained here.

The system in question is represented by the following non-linear differential equation

$$\ddot{x} + \dot{x} + x - \alpha x^3 = y(t). \qquad (3.132)$$

Christensen [27] has solved this equation via a Volterra series by considering it as representing a single input/single output non-linear system. However, in this section the equation will be solved by casting it into the form of (3.1) and utilizing the multidimensional Volterra series approach previously developed. The form of (3.1) will be obtained by methods stated in Ogata [30], that is:

(i) normal state form (series method)

(ii) Heavisides decompositon (parallel method)

3.4.1 Normal State Form

Equation (3.132) can be rewritten as

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = y(t) - x_1 - x_2 + \alpha x_1^3 ,$$

which can be written in vector-matrix form as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + |\alpha| \begin{bmatrix} 0 \\ 1 \end{bmatrix} x_1^3 + \begin{bmatrix} 0 \\ 1 \end{bmatrix} y(t). \tag{3.133}$$

Clearly (3.133) is of the same form as (3.1) the solution of which is given by equation (3.2).

The state transition matrix $\underline{\Phi}(t,s)$ of equation (3.2) is found from

$$\underline{\Phi}(t,s) = \mathcal{L}^{-1} (m\underline{I} - \underline{A})^{-1}, \tag{3.134}$$

where $\mathcal{L}^{-1}$ (----) is the inverse Laplace transform.

Therefore, from (3.133)

$$\underline{\Phi}(t,s) = \mathcal{L}^{-1} \begin{bmatrix} \dfrac{m+1}{m^2+m+1} & \dfrac{1}{m^2+m+1} \\ \dfrac{-1}{m^2+m+1} & \dfrac{}{m^2+m+1} \end{bmatrix}$$

$$= \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \tag{3.135}$$

where

$$\Phi_{11}(t,s) = \mathcal{L}^{-1} \frac{m + 1}{m^2 + m + 1}$$

$$= e^{-\frac{1}{2}(t-s)} \left(\cos \frac{\sqrt{3}}{2}(t-s) + \frac{1}{\sqrt{3}} \operatorname{Sin} \frac{\sqrt{3}}{2}(t-s)\right) \tag{3.136}$$

and

$$\Phi_{12}(t,s) = \mathcal{L}^{-1} \frac{1}{m^2 + m + 1}$$

$$= e^{-\frac{1}{2}(t-s)} \operatorname{Sin} \frac{\sqrt{3}}{2}(t-s) . \tag{3.137}$$

But $\Phi_{12}(t,s) = -\Phi_{21}(t,s)$ therefore,

$$\Phi_{21}(t,s) = -e^{-\frac{1}{2}(t-s)} \operatorname{Sin} \frac{\sqrt{3}}{2}(t-s) , \tag{3.138}$$

and finally

$$\Phi_{22}(t,s) = \mathcal{L}^{-1} \frac{m}{m^2 + m + 1}$$

$$= e^{-\frac{1}{2}(t-s)} \left(\operatorname{Cos} \frac{\sqrt{3}}{2}(t-s) - \frac{1}{\sqrt{3}} \operatorname{Sin} \frac{\sqrt{3}}{2}(t-s)\right). \tag{3.139}$$

It was shown in equation (3.43) that

$$P_{11} = \int_0^{t=\infty} |\Phi_{11}(t,s)| ds , \tag{3.140}$$

Similarly with P_{12}, P_{21}, P_{22}. Rewriting (3.136) in polar form results in

$$\Phi_{11}(t,s) = R\, e^{a(t-s)} \operatorname{Sin}(b(t-s) + \theta) \tag{3.141}$$

where

$$\begin{aligned} R &= 1{\cdot}15 \quad = 2/\sqrt{3} \\ b &= {\cdot}866 \quad = \sqrt{3}/2 \\ \theta &= 1{\cdot}05 \text{ radians} \\ a &= -{\cdot}5\ . \end{aligned} \tag{3.142}$$

Let

$$b(t-s) + \theta = bz \tag{3.143}$$

Then

$$t-s = z - \theta/b \tag{3.144}$$

now readjust the limits of integration to be in terms of z rather than s, then from (3.144), when

$$s = o \text{ then } z = t + \theta/b$$

therefore when

$$s = t = o \text{ then } z = \theta/b \tag{3.145}$$

and when

$$s = t \rightarrow \infty \text{ then } z \rightarrow \infty,$$

therefore the upper and lower limits are ∞ and θ/b respectively.

Substituting (3.143) and (3.144) into (3.141) results in

$$\Phi_{11}(t,s) = R\,e^{-a\theta/b}(e^{az}\text{Sin } bz). \tag{3.146}$$

Now substitute (3.146) into (3.140) and adjust the limits of integration in accordance with (3.145) with the result that (3.140) can now be written as

$$P_{11} = R\,e^{-a\theta/b}\int_{\theta/b}^{\infty}|e^{az}\text{Sin}\,bz|dz$$

$$= R\,e^{-a\theta/b}\left(\int_{0}^{\infty}e^{az}|\text{Sin}bz|dz - \int_{0}^{\theta/b}e^{az}|\text{Sin}\,bz|dz\right) \tag{3.147}$$

which may also be written as

$$P_{11} = R\,e^{-a\theta/b}(I_2 - I_1) \tag{3.148}$$

where I_1 and I_2 are the integrals in the brackets. The integral I_1 is evaluated first

$$I_1 = \int_{0}^{\theta/b}e^{az}|\,\text{Sin}bz|dz\ , \tag{3.149}$$

but since

$$o < \theta/b < \pi/b \tag{3.150}$$

The absolute value sign of (3.149) may be removed and integration performed directly, that is

$$I_1 = \int_0^{\theta/b} e^{az} \mathrm{Sin}\ bz\ da$$

$$= \frac{e^{az}}{a^2 + b^2} (a\mathrm{Sin}\ bz - b\mathrm{Cos}\ bz) \Big|_0^{\theta/b}$$

$$= \frac{e^{a\theta/b}}{a^2 + b^2} (a\mathrm{Sin}\ \theta - b\mathrm{Cos}\ \theta) + \frac{b}{a^2 + b^2} . \tag{3.151}$$

By substituting the values from (3.142) into equation (3.151), I_1 can be evaluated as

$$I_1 = \frac{e^{a\theta/b}}{a^2 + b^2} (a\mathrm{Sin}\ \theta - b\mathrm{Cos}\ \theta) + \frac{b}{a^2 + b^2}$$

$$= \cdot 395. \tag{3.152}$$

Now consider the second integral of (3.148) namely I_2

$$I_2 = \int_0^{\infty} e^{az} | \sin bz | \ dz \tag{3.153}$$

which is solved in the following manner.

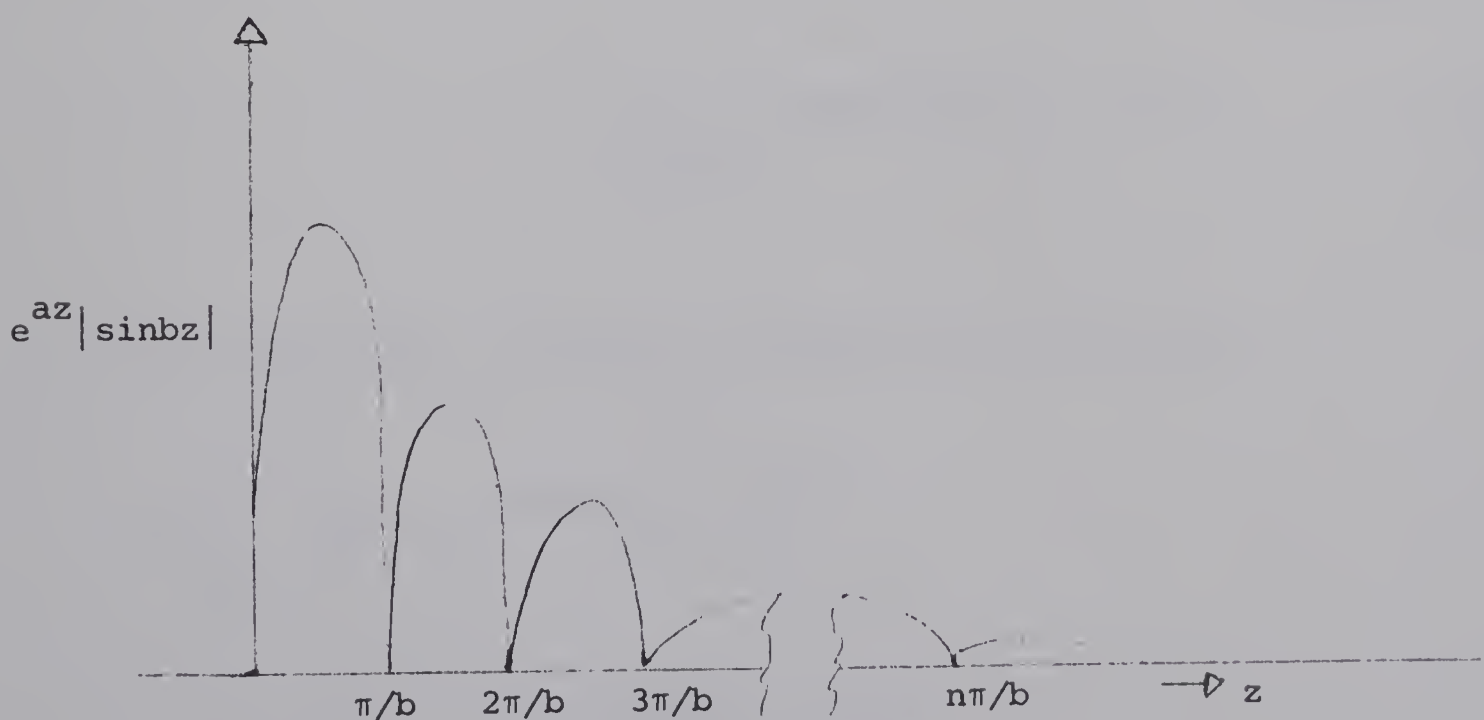

Figure 3: Relationship Between z and $e^{az}|\text{sinbz}|$.

From Figure 3

$$I_2 = \sum_{n=o}^{\infty} \int_{2n\pi/b}^{(2n+1)\pi/b} e^{az} \text{Sin } bz \, dz$$

$$- \sum_{n=o}^{\infty} \int_{(2n+1)\pi/b}^{(2n+2)\pi/b} e^{az} \sin bz \, dz \qquad (3.154)$$

To determine I_2 in (3.154) we find

$$\int_{2n\pi/b}^{(2n+1)\pi/b} e^{az} \text{Sin } bz \, dz = \frac{e^{az}}{a^2 + b^2} (a\text{Sin } bz - b\text{Cos } bz) \Big|_{2n\pi/b}^{(2n+1)\pi/b}$$

$$= \frac{b}{a^2 + b^2} e^{2na\pi/b} (1 + e^{a\pi/b}) \qquad (3.155)$$

and

$$\int_{(2n+1)\pi/b}^{(2n+2)\pi/b} e^{az} \text{Sin } bz \, dz = \frac{e^{az}}{a^2 - b^2} (a\sin bz - b\text{Cos } bz) \Big|_{(2n+1)\pi/b}^{(2n+2)\pi/b}$$

$$= \frac{-b}{a^2 + b^2} e^{(2n+1)a\pi/b} (1 + e^{a\pi/b}). \qquad (3.156)$$

If we now substitute (3.155) and (3.156) into (3.154) then

$$I_2 = \sum_{n=o}^{\infty} \frac{b}{a^2 + b^2} e^{2na\pi/b} (1 + e^{a\pi/b})$$

$$+ \sum_{n=o}^{\infty} \frac{b}{a^2 + b^2} e^{(2n+1)a\pi/b} (1 + e^{a\pi/b})$$

$$= \frac{b}{a^2 + b^2} (1 + e^{a\pi/b}) \sum_{n=o}^{\infty} (e^{2na\pi/b} + e^{(2n+1)a\pi/b})$$

$$= \frac{b}{a^2 + b^2} \left(1 + e^{a\pi/b}\right)^2 \sum_{n=o}^{\infty} e^{2na\pi/b}. \tag{3.157}$$

For convenience let

$$r = e^{2a\pi/b} \tag{3.158}$$

thus

$$\sum_{n=o}^{\infty} r^n = \sum_{n=o}^{\infty} e^{2na\pi/b}. \tag{3.159}$$

Clearly (3.159) is a geometric series and the partial sum to "n" terms is

$$S_n = \frac{1 - r^{n+1}}{1 - r}, \tag{3.160}$$

and from the values given in (3.142) it is found that

$$a<o \tag{3.161}$$

therefore, because $r = e^{2a\pi/b}$ then

$$r<1, \tag{3.162}$$

thus as $n\to\infty$, $r^{n+1}\to o$, thus

$$S_{n\to\infty} = \frac{1}{1 - r} \ . \tag{3.163}$$

Hence substituting (3.158) into (3.163) results in

$$S_{n\to\infty} = \frac{1}{1 - e^{2a\pi/b}}$$

$$= \sum_{n=o}^{\infty} e^{2na\pi/b}. \tag{3.164}$$

Now substitute (3.164) into (3.157) and write

$$I_2 = \frac{b}{a^2 + b^2} (1 + e^{a\pi/b})^2 \frac{1}{1 - e^{2a\pi/b}}$$

$$= \frac{b}{a^2 + b^2} \left(\frac{1 + e^{a\pi/b}}{1 - e^{a\pi/b}}\right). \tag{3.165}$$

Then by substituting values from (3.142) into (3.165) I_2 can be evaluated as

$$I_2 = \frac{\cdot 866}{\cdot 5^2 + \cdot 866^2} \left(\frac{1 + e^{-\cdot 5\pi/\cdot 866}}{1 - e^{-\cdot 5\pi/\cdot 866}}\right) \tag{3.166}$$

$$= \cdot 866 \left(\frac{1\cdot 16}{\cdot 838}\right)$$

$$= 1\cdot 20 \ .$$

Substituting (3.166) and (3.152) into (3.149) results in

$$P_{11} = Re^{-a\theta/b}(1{\cdot}20 - {\cdot}395)$$
$$= Re^{-a\theta/b}({\cdot}805) \ , \tag{3.167}$$

and if values from (3.142) are now substituted into (3.167), then P_{11} is found to be

$$P_{11} = 1{\cdot}15 \ \ x \ {\cdot}805 \ e^{-{\cdot}5x1{\cdot}05/{\cdot}866}$$
$$= 1{\cdot}71 \ . \tag{3.168}$$

Similarly with P_{12} where

$$P_{12} = \int_0^\infty |\Phi_{12}(t,s)|ds. \tag{3.169}$$

Now rewrite (3.137) in polar form and by utilizing the values given in (3.142), equation (3.169) can be cast into a form similar to (3.147). Thus (3.169) may be rewritten as

$$P_{12} = \int_0^\infty e^{az}\left|\text{Sin}\, bz\right|dz \tag{3.170}$$

which is clearly of the same form as (3.153) the solution of which is (3.166); therefore

$$P_{12} = 1{\cdot}20. \tag{3.171}$$

By comparing (3.137) and (3.138) it is seen that

$$P_{12} = -P_{21} \tag{3.172}$$

therefore,

$$P_{21} = - 1{\cdot}20. \tag{3.173}$$

The last element of the $\underline{P}$ matrix, namely P_{22}, is determined in a manner identical to that used in the determination of P_{11}. That is,

$$\Phi_{22} = Re^{a\theta/b}(e^{az}\mathrm{Sin}\,bz). \tag{3.174}$$

therefore,

$$P_{22} = Re^{a\theta/b}\int_{-\theta/b}^{\infty}|e^{az}\mathrm{Sin}\,bz|\,dz$$

$$= Re^{a\theta/b}\left(\int_0^{\infty} e^{az}|\mathrm{Sin}\,bz|\,dz - \int_0^{-\theta/b} e^{az}|\mathrm{Sin}\,bz|\,dz\right)$$

$$= Re^{a\theta/b}(I_2 - I_1). \tag{3.175}$$

Evaluating I_1

$$I_1 = \frac{e^{az}}{a^2 + b^2}\,(a\mathrm{Sin}\,bz - b\mathrm{Cos}\,bz)\Big|_0^{-\theta/b} \tag{3.176}$$

or

$$I_1 = \frac{-e^{-a\theta/b}}{a^2 + b^2}\,(a\mathrm{Sin}\,\theta + b\mathrm{Cos}\,\theta) + \frac{b}{a^2 + b^2}\,.$$

By direct substitution of values from (3.142) into equation (3.176), it is found that

$$I_1 = -1{\cdot}83\ (-{\cdot}5\ \mathrm{Sin}(1{\cdot}05) + {\cdot}866\ \mathrm{Cos}(1{\cdot}05)) + {\cdot}866$$

$$= {\cdot}866\ , \qquad (3.177)$$

I_2 can be evaluated by utilizing equation (3.165) which is rewritten here for convenience

$$I_2 = \frac{b}{a^2 + b^2}\left(\frac{1 + e^{a\pi/b}}{1 - e^{a\pi/b}}\right),$$

and in accordance with equation (3.166) it is found that

$$I_2 = 1{\cdot}20\ . \qquad (3.178)$$

Now substitute (3.178) and (3.177) into (3.175), and write

$$P_{22} = Re^{a\theta/b}(1{\cdot}20 - {\cdot}866) \qquad (3.179)$$

then by substituting values from (3.142) into (3.179) P_{22} is found to be

$$P_{22} = 1{\cdot}15\ \times\ {\cdot}334e^{-{\cdot}5\ \times\ 1{\cdot}05/{\cdot}866}$$

$$= {\cdot}209\ . \qquad (3.180)$$

The elements of the $\underline{P}$ matrix have now been evaluated, thus

$$\underline{P} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \tag{3.181}$$

$$= \begin{bmatrix} 1\cdot71 & 1\cdot20 \\ -1\cdot20 & \cdot209 \end{bmatrix} \tag{3.182}$$

where values of the elements of $\underline{P}$ are given by (3.168), (3.171), (3.173) and (3.180).

The "size" of the vectors which satisfy the contraction mapping and fixed point condition can now be determined. The contraction condition (3.55) **requires** that the norm of $\underline{Pz}(t)$ be less than one. The $\underline{P}$ matrix has already been found (3.182), now $\underline{z}(t)$ will be determined from (3.44) and (3.133); therefore,

$$\underline{z} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} = \sup_t |\alpha| \begin{bmatrix} o & o \\ 3x_1^2 & o \end{bmatrix} . \tag{3.183}$$

Thus the resulting matrix product of equation (3.47) is

$$\left\|\underline{P}\ \underline{z}\right\| \leq |\alpha| \ \sup \left| \begin{bmatrix} 1\cdot71 & 1\cdot20 \\ -1\cdot20 & \cdot209 \end{bmatrix} \begin{bmatrix} o & o \\ 3x_1^2 & o \end{bmatrix} \right| \tag{3.184}$$

which reduces to

$$\left\|\underline{P}\ \underline{z}\right\| \leq |\alpha| \ 3\cdot60\underline{X}^{*2} \tag{3.185}$$

where

$$\underline{\bar{X}}^* \underline{<} \sup_i \sup_t \left| x_i(t) \right|, \tag{3.186}$$

and from (3.55) it follows that

$$|\alpha| 3 \cdot 60 \underline{\bar{X}}^{*2} \underline{<} K < 1. \tag{3.187}$$

Similarly with the fixed point condition, equation (3.102), the vector $\underline{L}$ may be obtained by substituting elements of (3.133) into (3.98); that is

$$\underline{L} = \begin{bmatrix} L_1 \\ \\ L_2 \end{bmatrix} = \sup_t |\alpha| \begin{bmatrix} o \\ \\ x_1^{\ 3} \end{bmatrix} . \tag{3.188}$$

By utilizing the $\underline{P}$ matrix (3.182) we find that

$$\|\underline{P}\ \underline{L}\| \underline{<} |\alpha| \ \sup \left| \begin{bmatrix} 1 \cdot 71 & 1 \cdot 20 \\ \\ -1 \cdot 20 & \cdot 209 \end{bmatrix} \begin{bmatrix} o \\ \\ x_1^{\ 3} \end{bmatrix} \right| \tag{3.189}$$

which reduces to

$$\|\underline{P}\ \underline{L}\| \underline{<} |\alpha| \ 1 \cdot 20 \underline{\bar{X}}^{*3} \tag{3.190}$$

(see equation (3.186) for norm on $\underline{\bar{X}}^*$).

Substitute (3.190) into (3.102) and write

$$|\alpha| \ 1 \cdot 20 \underline{\bar{X}}^{*3} < (1 - K)\ U\,. \tag{3.191}$$

And from (3.187) $\underline{\bar{X}}^*$ can be evaluated as

$$\underline{\bar{X}}^* \leq \left(\frac{K}{3 \cdot 60 |\alpha|} \right)^{\frac{1}{2}} \tag{3.192}$$

and from (3.191)

$$\underline{\bar{X}}^* < \left(\frac{(1 - K)U}{1 \cdot 20 |\alpha|} \right)^{1/3} , \tag{3.193}$$

now let

$$K = \cdot 6$$

$$U = \cdot 2$$

$$\alpha = 1 \ .$$

Then from (3.192)

$$\underline{\bar{X}}^* \leq \cdot 408 \tag{3.194}$$

and from (3.193)

$$\underline{\bar{X}}^* < \cdot 405 \tag{3.195}$$

taking the supremum of (3.194) and (3.195) it is found that,

$$\underline{\bar{X}}^* < \cdot 405 \ . \tag{3.196}$$

Hence a bound on the output vector has been determined, via a Volterra series.

3.4.2 Heavisides Decomposition

It is generally recognized that the definition of state variable relations is not unique. The state variable relations in this section will be obtained via Heavisides decomposition.

Equation (3.132) can be represented in the following diagramatic form

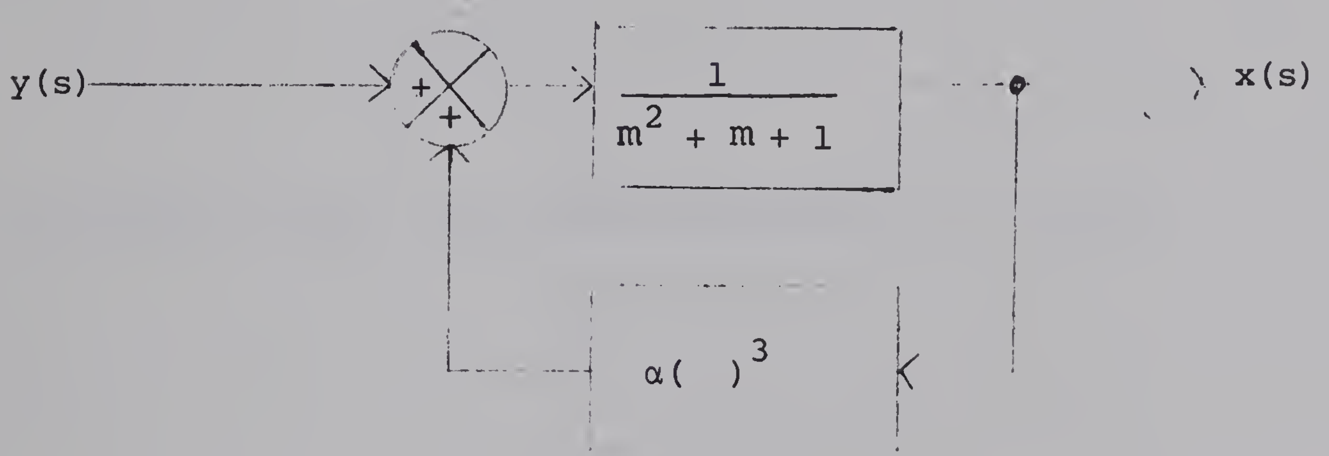

Figure 4: Block Diagram of Equation (3.132)

Furthermore $\dfrac{1}{m^2 + m + 1}$ can be decomposed as

$$\frac{1}{m^2 + m + 1} = \frac{C_1}{m + \lambda_1} + \frac{C_2}{m + \lambda_2} \tag{3.197}$$

where

$$\lambda_1 = \frac{-1}{2} + j\,\frac{\sqrt{3}}{2} \tag{3.198}$$

and

$$\lambda_2 = \frac{-1}{2} - j\frac{\sqrt{3}}{2} , \tag{3.199}$$

therefore,

$$C_1 = \frac{j}{\sqrt{3}} \tag{3.200}$$

and

$$C_2 = \frac{-j}{\sqrt{3}} \quad . \tag{3.201}$$

The system of Fig. 4 can now be represented as follows.

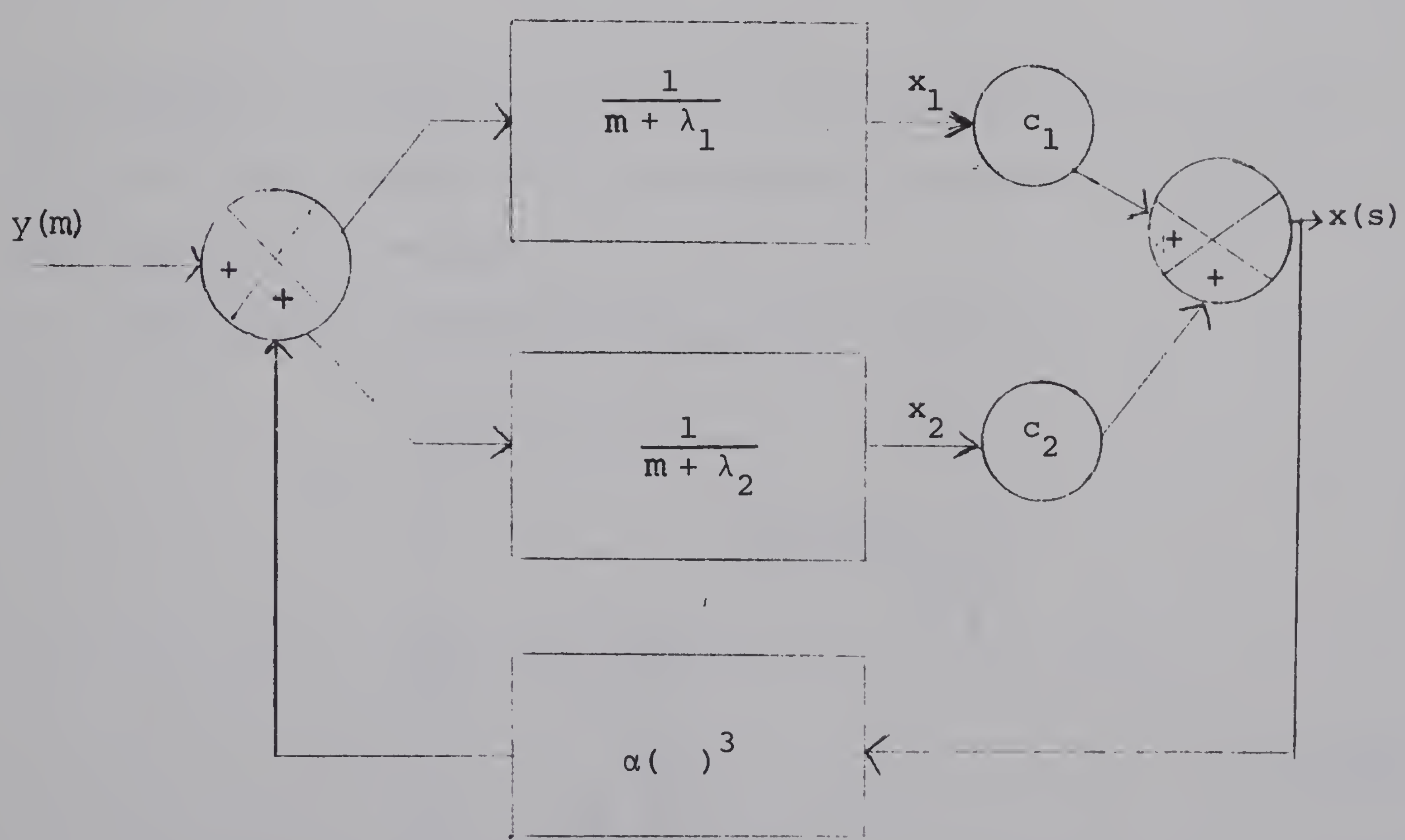

Figure 5: Parallel Form of Equation (3.132)

where $\underline{x} = c_1x_1 + c_2x_2$. (3.202)

From Fig. 5 it is seen that

$$\dot{x}_1 = y(t) - \alpha(x)^3 - \lambda_1 x_1$$
$$\dot{x}_2 = y(t) - \alpha(x)^3 - \lambda_2 x_2. \tag{3.203}$$

Now substitute (3.202) into (3.203), and write the resulting equations in matrix form as follows:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -\lambda_1 & 0 \\ 0 & -\lambda_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} -\alpha(c_1x_1 + c_2x_2)^3 \\ -\alpha(c_1x_1 + c_2x_2)^3 \end{bmatrix} + \begin{bmatrix} y(t) \\ y(t) \end{bmatrix}. \tag{3.204}$$

By comparing (3.204) and (3.133) it is seen that they are of the same form, thus the procedure which was utilized in section 3.4.1 will be used again in this section.

The $\underline{P}$ matrix is determined first and is found to be

$$\underline{P} = \begin{bmatrix} \int_0^\infty |e^{-\lambda_1 t}|dt & 0 \\ 0 & \int_0^\infty |e^{-\lambda_2 t}|dt \end{bmatrix}$$
$$= \begin{bmatrix} \frac{1}{\lambda_1} & 0 \\ 0 & \frac{1}{\lambda_2} \end{bmatrix} . \tag{3.205}$$

The elements of the $\underline{z}$ matrix are constructed from (3.204) via equation (3.45), typically

$$z_{kj} = |\alpha| \sup_t \left|\frac{\partial f_k}{\partial x_j}\right| \tag{3.206}$$

therefore,

$$z_1 = |\alpha| \sup_t (|c_1| \left|\frac{\partial f_1}{\partial x_1}\right| + |c_2| \left|\frac{\partial f_1}{\partial x_2}\right| \qquad o) \tag{3.207}$$

which differs from Z_2 only in the subscript attached to "∂f", and

$$|c_1| = |c_2| = \frac{1}{\sqrt{3}}. \tag{3.208}$$

also

$$\left|\frac{\partial f_1}{\partial x_1}\right| = \left|\frac{\partial f_1}{\partial x_2}\right| = \left|3(c_1x_1 + c_2x_2)^2\right|. \tag{3.209}$$

Thus,

$$z_1 = z_2 \leq (|\alpha| 3 \left|(c_1x_1 + c_2x_2)^2\right| \Big((|c_1| + |c_2|) \qquad o\Big)) \tag{3.210}$$

$$\leq (\frac{6|\alpha|}{\sqrt{3}} |\underline{x}|^2 \qquad o). \tag{3.211}$$

The matrix product of equation (3.47) can now be written as

$$\sup_t \left\| \underline{P}\ \underline{z} \right\|$$

$$\leq |\alpha| \sup_i \sup_t \left| \begin{bmatrix} \frac{1}{\lambda_1} & 0 \\ 0 & \frac{1}{\lambda_2} \end{bmatrix} \begin{bmatrix} \frac{6|\alpha|\ |\underline{x}|^2}{\sqrt{3}} & 0 \\ \frac{6|\alpha|\ |\underline{x}|^2}{\sqrt{3}} & 0 \end{bmatrix} \right| \tag{3.212}$$

and because

$$|\lambda_1| = |\lambda_1| = 1 \tag{3.213}$$

then (3.212) reduces to

$$\sup_t \left\| \underline{P}\ \underline{z} \right\| \leq \frac{6|\alpha|\underline{\bar{X}}^{*2}}{\sqrt{3}} \tag{3.214}$$

substitute (3.214) into (3.55) and write

$$\frac{6|\alpha|\underline{\bar{X}}^{*2}}{\sqrt{3}} \leq K<1 \tag{3.215}$$

therefore,

$$\underline{\bar{X}}^{*} \leq \left(\frac{K}{3 \cdot 46|\alpha|}\right)^{\frac{1}{2}}. \tag{3.216}$$

The elements of the $\underline{L}$ vector can be constructed from (3.213) via (3.98), typically

$$L_k \leq |\alpha| \sup_t \left| f_k(\underline{x}(t)) \right| \tag{3.217}$$

therefore

$$L_1 = L_2 \leq |\alpha| \sup_t \left| (c_1 x_1 + c_2 x_2)^3 \right| \tag{3.218}$$

$$\leq |\alpha| \sup_t \left| \underline{x} \right|^3 .$$

Thus the matrix vector product of (3.102) can now be written as

$$|\alpha| \sup_i \sup_t \left| \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} |\underline{x}|_3^3 \\ |\underline{x}| \end{bmatrix} \right| < (1 - K)\,U \tag{3.219}$$

which reduces to

$$|\alpha| \underline{\overline{X}}^{*3} < (1 - K)\,U \tag{3.220}$$

therefore,

$$\underline{\overline{X}}^{*} < \left(\frac{(1 - K)U}{|\alpha|}\right)^{1/3}, \tag{3.221}$$

now let

$$K = \cdot 6$$
$$U = \cdot 2$$
$$\alpha = 1 .$$

Then from (3.216)

$$\underline{\overline{X}}^{*} \leq \cdot 416 \tag{3.222}$$

and from (3.221)

$$\underline{\overline{X}}^{*} \underline{\leq} \cdot 430, \tag{3.223}$$

taking the supremum of (3.222) and (3.223) it is found that

$$\underline{\overline{X}}^{*} \underline{\leq} \cdot 416 \tag{3.224}$$

Hence a bound on the output vector has been determined by the Volterra series method.

The bound on the input vector $\underline{y}(t)$ will now be determined. Consider equation (3.5) once again, if the initial conditions are allowed to be zero then the equation may be written as

$$\underline{x}^{*} = \underline{\ell} + \underline{u}^{*}. \tag{3.225}$$

Take the norm of (3.225) and write

$$\underline{\overline{X}}^{*} \underline{\leq} \mathcal{L} + U , \tag{3.226}$$

from the material in sections 3.1 and 3.2 it can be clearly seen that

$$\mathcal{L} \underline{\leq} \sup|\underline{P}\ \underline{y}|. \tag{3.227}$$

Thus $\mathcal{L}$ can be determined from the $\underline{P}$ matrix (3.182) and the $\underline{y}$ vector (3.133), that is

$$\mathcal{L} \leq \sup \left| \begin{bmatrix} 1{\cdot}71 & 1{\cdot}20 \\ -1{\cdot}20 & {\cdot}209 \end{bmatrix} \begin{bmatrix} o \\ y \end{bmatrix} \right| \tag{3.228}$$

$$\leq 1{\cdot}20\overline{y} \tag{3.229}$$

where

$$\overline{Y} \leq \sup_i \sup_t \left| y_i(t) \right| . \tag{3.230}$$

From (3.229) $\mathcal{L}$ can be rewritten as

$$\mathcal{L} \leq H\overline{Y} \tag{3.231}$$

where

$$H = 1{\cdot}20, \tag{3.232}$$

then (3.226) can be rewritten as

$$\overline{X}^* \leq H\overline{Y} + U . \tag{3.233}$$

Substitute (3.233) into the contraction condition (3.192) and write

$$H\overline{Y} + U \leq \left(\frac{K}{3|\alpha|H}\right)^{\frac{1}{2}} \tag{3.234}$$

and by substituting the values assumed for K, α, and U into (3.234) $\overline{Y}$ is found to be

$$\overline{Y} \leq {\cdot}173 . \tag{3.235}$$

From the fixed point condition (3.93), and equation (3.15) is is easily shown that

$$H\overline{\underline{Y}} < \left(\frac{(1 - K)\ U}{H|\alpha|}\right)^{1/3} \qquad (3.236)$$

and by substituting the values assumed for K, α, and U into (3.236), $\overline{\underline{Y}}$ is found to be

$$\overline{\underline{Y}} < \cdot 375\ . \qquad (3.237)$$

Taking the supremum of (3.235) and (3.237) it is found that

$$\overline{\underline{Y}} \leq \cdot 173\ . \qquad (3.238)$$

Hence a bound on the input vector has been determined.

Therefore, by the contraction fixed point principle

$$\begin{aligned} &\overline{\underline{Y}} \leq \cdot 173 \\ &\overline{\underline{X}}^{*} < \cdot 405\ . \end{aligned} \qquad (3.239)$$

<u>Conclusion</u>:

By normal state matrix methods and the Volterra series the system considered has been shown to be B.I.B.O. stable.

Now determine the bound on the input vector for the parallel system i.e. section 3.4.2. Thus $\mathcal{L}$ can be determined from the $\underline{P}$ matrix (3.205)

and (3.213) and the $\underline{Y}$ vector (3.204), that is

$$\mathcal{L} \leq \sup \left| \begin{bmatrix} 1 & o \\ o & 1 \end{bmatrix} \begin{bmatrix} y \\ y \end{bmatrix} \right|$$

$$\leq \underline{\bar{Y}} \tag{3.240}$$

where $\underline{\bar{Y}}$ is defined by (3.230)

and

$$H = 1 \ . \tag{3.241}$$

If we now substitute (3.241) into (3.234) then

$$\underline{\bar{Y}} \leq \cdot 226, \tag{3.242}$$

also if we substitute (3.241) into (3.236) then

$$\underline{\bar{Y}} < \cdot 430 \tag{3.243}$$

and since

$$\cdot 226 < \cdot 430$$

This implies that

$$\underline{\bar{Y}} \leq \cdot 226 \ . \tag{3.244}$$

Therefore by the contraction fixed point principle

$$\underline{\bar{Y}} \underline{<} \cdot 226$$

$$\underline{\bar{X}}^{*} \underline{<} \cdot 416 \qquad (3.245)$$

<u>Conclusion</u>:

The parallel method of erecting the state equations and the subsequent use of the Volterra series has shown that B.I.B.O. stability exists.

In the worked example the $\underline{A}$ matrix was a constant, however, if $\underline{A}$ is time dependent then the commuting rule which is invoked in linear systems must be invoked here, appropriately, since a linear system is a special case of a nonlinear system.

CHAPTER IV

SAMPLED DATA CONTROL SYSTEMS

4.0 Introduction

A very important sub-class of control systems are those of the sampled data type. Sampled data control systems consist of dynamic systems whose inputs are either time sequences or are important only at specific instants of time.

The two types of sampled data systems considered here are defined as,

Type "1":

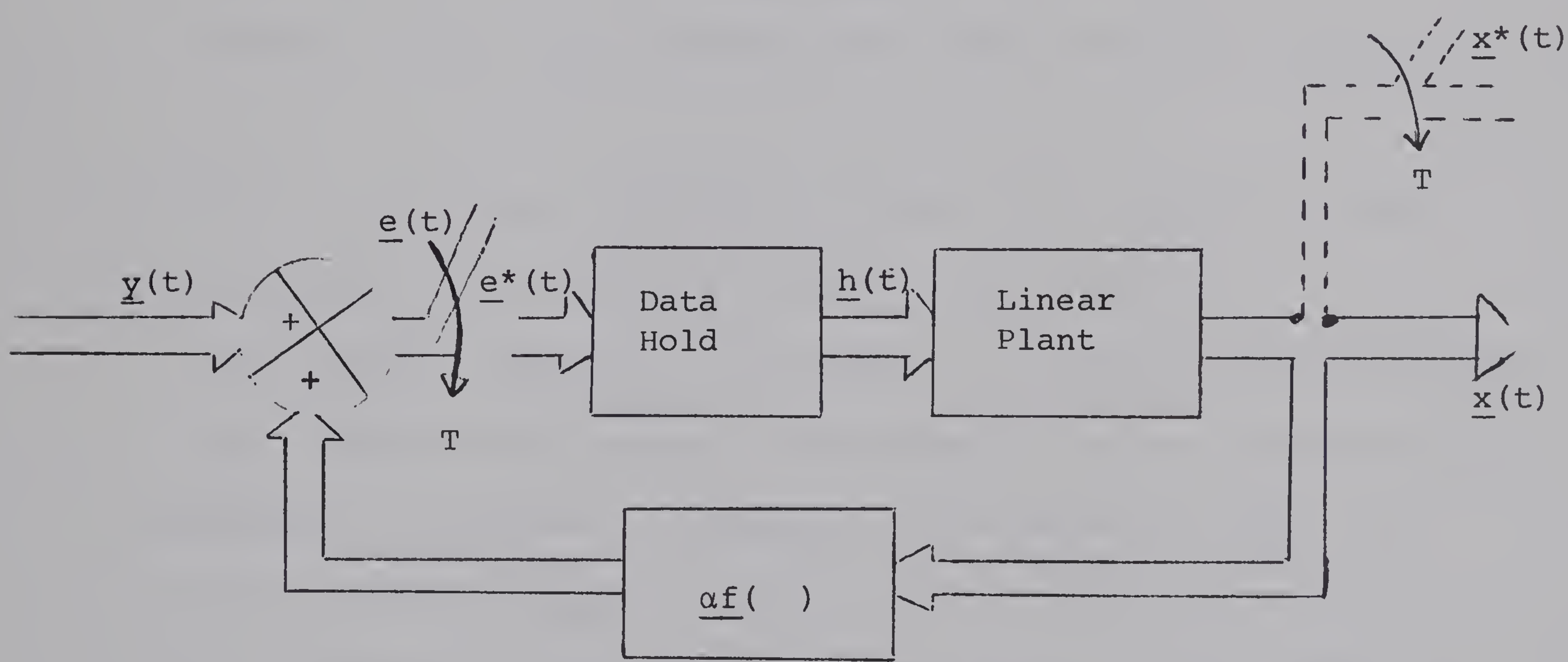

Figure 6: Non-linear sampled data system with data hold.

The error signal $\underline{e}(t)$ is sampled and fed into the data hold which maintains a steady input to the linear plant during the in-between sample periods.

Type "2":

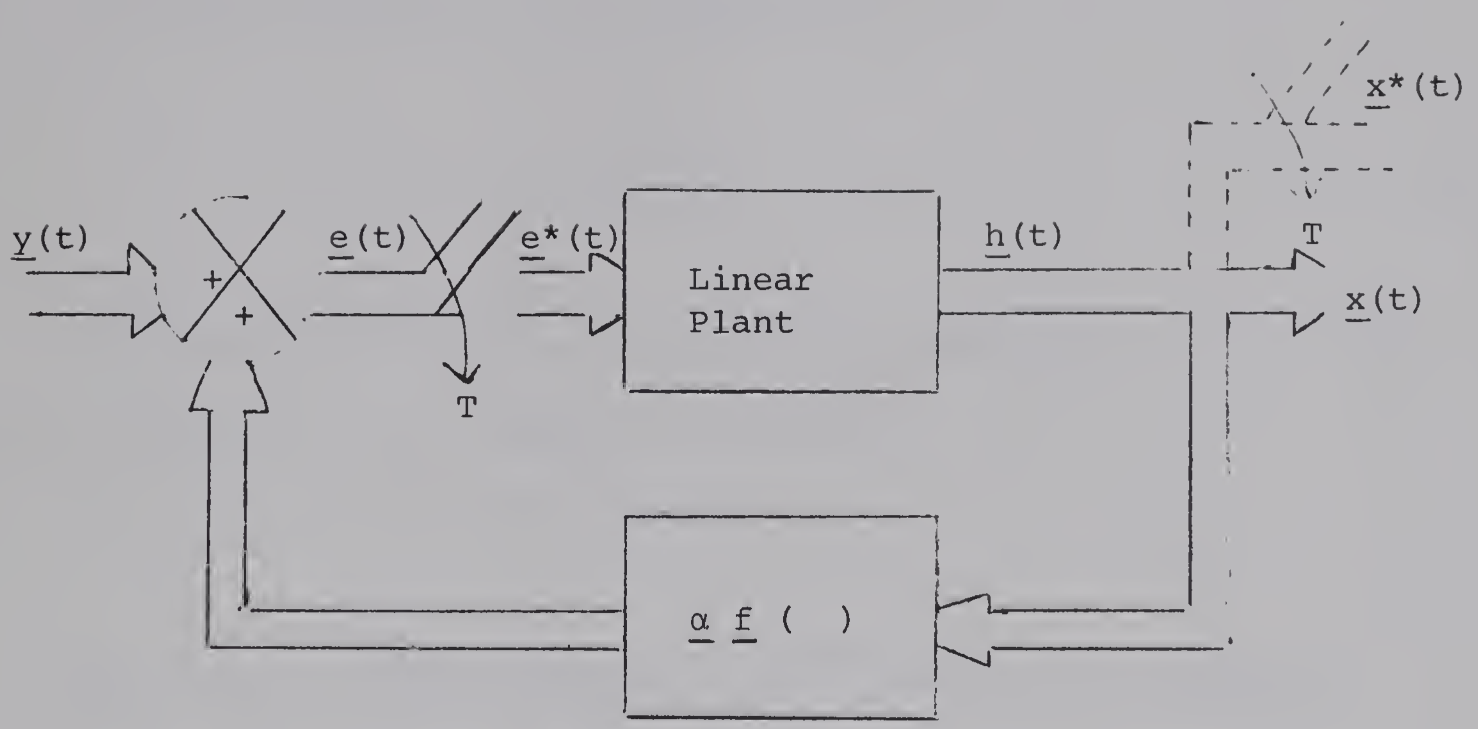

Figure 7: Non-linear sampled data system without data hold.

The input to the linear plant $\underline{e}^*(t)$, appears as a series of pulses. Thus the plant has an input only during sample periods.

The material developed in the preceeding sections will be extended to cover discrete data systems. The components of the systems are still describable by differential equations, but because the signal is discrete a set of difference equations will be generated from the original differential equations. It is well known, [16], [30],[32], that difference equations represent discrete systems in a manner analogous to the way differential equations represent continuous systems.

4.1 Determination of Region of Stability for a Type "1" System

Consider equation (3.1) which represents a continuous non-linear system, it is repeated here for convenience

$$\dot{\underline{x}}(t) = \underline{A}\underline{x}(t) + \underline{B}\underline{y}(t) + \underline{\alpha}\ \underline{f}(x(t)). \qquad (4.1)$$

The solution of (4.1) is well known,

$$\underline{x}(t) = \underline{\Phi}(t)\underline{x}(o) + \int_0^t \underline{\Phi}(t,s)\underline{By}(s)ds$$

$$+ \int_0^t \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}(s))ds. \tag{4.2}$$

A method of descretizing (4.2) has been developed by Ogata, [30], it is repeated here for reasons of continuity.

Let t = (n + 1)T in (4.2), thus

$$\underline{x}((n+1)T) = \underline{\Phi}((n+1)T)\underline{x}(o) + \int_0^{(n+1)T} ((n+1)T,s)\ \underline{B}\ \underline{y}(s)ds$$

$$+ \int_0^{(n+1)T} \underline{\Phi}((n+1)T,s)\ \underline{\alpha}\ \underline{f}(\underline{x}(s))ds; \tag{4.3}$$

Let t = nT in (4.2) then

$$\underline{x}(nT) = \underline{\Phi}(nT)\underline{x}(o) + \int_0^{nT} \underline{\Phi}(nT,s)\ \underline{B}\ \underline{y}(s)ds$$

$$+ \int_0^{nT} \underline{\Phi}(nT,s)\ \underline{\alpha}\ \underline{f}(\underline{x}(s))ds. \tag{4.4}$$

Premultiply (4.4) by $\underline{\Phi}(T)$

$$\underline{\Phi}(T)\underline{x}(nT) = \underline{\Phi}((n+1)T)\underline{x}(o) + \int_0^{nT} \underline{\Phi}((n+1)T,s)\ \underline{B}\ \underline{y}(s)ds$$

$$+ \int_0^{nT} \underline{\Phi}((n+1)T,s)\ \underline{\alpha}\ \underline{f}(\underline{x}(s))ds. \tag{4.5}$$

Now subtract (4.5) from (4.3)

$$\underline{X}((n+1)T) - \underline{\Phi}(T)\underline{x}(nT) = \int_{nT}^{(n+1)T} ((n+1)T,s)\ \underline{B}\ \underline{y}(s)ds$$

$$+ \int_{nT}^{(n+1)T} \underline{\Phi}((n+1)T,s)\ \underline{\alpha}\ \underline{f}(\underline{x}(s))ds; \qquad (4.6)$$

Therefore,

$$\underline{x}((n+1)T) = \underline{\Phi}(T)\underline{x}(nT) + \int_{nT}^{(n+1)T} \underline{\Phi}((n+1)T,s)\ \underline{B}\ \underline{y}(s)ds$$

$$+ \int_{nT}^{(n+1)T} \underline{\Phi}((n+1)T,s)\ \underline{\alpha}\ \underline{f}(\underline{\mathbf{x}}(s))ds. \qquad (4.7)$$

Now let

$$-\tau = nT-s \qquad (4.8)$$

Therefore,

$$\underline{\Phi}((n+1)T,s) = \underline{\Phi}(T,\tau) \qquad (4.9)$$

and

$$d\tau = ds\ . \qquad (4.10)$$

The upper and lower limits of integration change from (n+1)T and nT to T and O respectively. Thus

$$\int_{nT}^{(n+1)T} \underline{\Phi}((n+1)T,s)\text{---}ds = \int_{0}^{T} \underline{\Phi}(T,\tau)\text{--}d\tau \qquad (4.11)$$

this follows from the preceeding remarks; and by making the necessary substitutions into (4.7). Equation (4.7) can now be written as,

$$\underline{x}((n+1)T) = \underline{\Phi}(T)\underline{x}(nT) + \int_0^T \underline{\Phi}(T,\tau)\underline{B}\ \underline{y}(\tau + nT)d\tau$$

$$+ \int_0^T \underline{\Phi}(T,\tau)\ \underline{\alpha}\ \underline{f}(\underline{x}(\tau + nT))d\tau\ . \tag{4.12}$$

But

$$\underline{h}(\tau) = \underline{h}(nT) \quad n = 0,1,2--- \tag{4.13}$$

where

$$nT \leq \tau < (n+1)T. \tag{4.14}$$

Therefore,

$$\underline{y}(\tau + nT) = \underline{y}(nT) \tag{4.15}$$

$$\underline{f}(\underline{x}(\tau + nT)) = \underline{f}(\underline{x}(nT)). \tag{4.16}$$

Since $\underline{y}(nT)$ and $\underline{f}(\underline{x}(nT))$ are constant between sampling instants they may be taken outside the integrals, thus

$$\underline{x}((n+1)T) = \underline{\Phi}(T)\underline{x}(nT) + \int_0^T \underline{\Phi}(T,\tau)\underline{B}d\tau\ \underline{y}(nT)$$

$$+ \int_0^T \underline{\Phi}(T,\tau)\ \underline{\alpha}\ d\tau\ \underline{f}(\underline{x}(nT)). \tag{4.17}$$

Equation (4.17) can be considered as representing an updated system at the start of each sampling instant. The bound on $\underline{x}((n+1)T)$ can be determined from equation (4.17) for some "n", by utilizing the methods of Chapter 3. However, the bound so determined is strictly local to that sample. What is actually required is the bound on the output over the range of samples, for $n \geq o$. Therefore, on this premise equation (4.17) must be modified such that $\underline{x}((n+1)T)$ is related directly to $\underline{x}(0)$, meaning, the initial conditions at n=0. Thus a table of $\underline{x}((n+1)T)$ for n=0,1,2---,m-1, will be constructed from which the general form of $\underline{x}((n+1)T) = \underline{x}(mT)$ can be found [32].

Because $\underline{x}((n+1)T)$ is a sampled output a superscript, a star(*), will be added to equation (4.17) to indicate sampled functions. Thus (4.17) is rewritten as

$$\underline{x}^*(mT) = \underline{x}^*((n+1)T) = \underline{\Phi}(T)\underline{x}^*(n\tau) + \int_0^T \underline{\Phi}(t,\tau)\ \underline{B}d\tau\ y^*(nT) + \int_0^T \underline{\Phi}(T,\tau)\ \underline{\alpha}\ d\tau\ \underline{f}(\underline{x}^*(nT)). \tag{4.19}$$

If $\underline{\alpha} = \underline{\theta}$ (the null matrix) then

$$\underline{x}^*(mT) = \underline{x}^*((n+1)T) = \underline{\Phi}(T)\underline{x}^*(nT) + \int_0^T \underline{\Phi}(T,\tau)\underline{B}\ d\tau\ \underline{y}^*(nT). \tag{4.20}$$

Which is the discrete equation for a linear system.

Now let n=0 in equation (4.19) then

$$\underline{x}^*(T) = \underline{\Phi}(T)\underline{x}^*(0) + \int_0^T \underline{\Phi}(T,\tau)\ \underline{B}\ d\tau\ \underline{y}^*(0)$$

$$+ \int_0^T \underline{\Phi}(T,\tau)\ \underline{\alpha}\ d\tau\ \underline{f}(\underline{x}^*(0)). \tag{4.21}$$

Let n=1 in (4.19)

$$\underline{x}^*(2T) = \underline{\Phi}(T)\underline{x}^*(T) + \int_0^T \underline{\Phi}(T,\tau)\ \underline{B}\ d\tau\ \underline{y}^*(T)$$

$$+ \int_0^T \underline{\Phi}(T,\tau)\ \underline{\alpha}\ d\tau\ \underline{f}(\underline{x}^*(T)). \tag{4.22}$$

Now substitute (4.21) into (4.22)

$$\underline{x}^*(2T) = \underline{\Phi}^2(T)\underline{x}^*(0) + \underline{\Phi}(T)\int_0^T \underline{\Phi}(T,\tau)\ \underline{B}\ d\tau\ \underline{y}^*(0)$$

$$+ \int_0^T \underline{\Phi}(T,\tau)\ \underline{B}\ d\tau\ \underline{y}^*(T) + \underline{\Phi}(T)\int_0^T \underline{\Phi}(T,\tau)\ \underline{\alpha}\ d\tau\ \underline{f}(\underline{x}^*(0))$$

$$+ \int_0^T \underline{\Phi}(T,\tau)\ \underline{\alpha}\ d\tau\ \underline{f}(\underline{x}^*(T)) \tag{4.23}$$

and so on; until after "m" terms.

$$\underline{x}^*(mT) = \underline{\Phi}^m(T)\underline{x}^*(0) + \sum_{i=o}^{m-1} \underline{\Phi}^{m-1-i}(T)\int_0^T \underline{\Phi}(T,\tau)\ \underline{B}\ d\tau\ \underline{y}^*(iT)$$

$$+ \sum_{i=o}^{m-1} \underline{\Phi}^{m-1-i}(T)\int_0^T \underline{\Phi}(T,\tau)\ \underline{\alpha}\ d\tau\ \underline{f}(\underline{x}^*(iT)). \tag{4.24}$$

The discrete state transition equation (4.24) is analogous to, and a special case of, its continuous counter part, equation (4.2). Therefore, (4.24) can be solved by methods analogous to those used in the solution of (4.2). The necessary equations will be suitably modified to accommodate a discrete data system fitted with a data hold device.

It is seen that the output $\underline{x}^*(mT)$ is the summation of all the past contributions, resulting from input pulses ranging from t=0 to t=mT. The upper limit of the summation in (4.24) can be extended to infinity because $\underline{\Phi}((m-1-i)T) = 0$ if i>m-1. Therefore (4.24) can be written as

$$\underline{x}^*(mT) = \underline{\Phi}^m(T)\underline{x}^*(0) + \sum_{i=o}^{\infty} \underline{\Phi}^i(T)\int_0^T \underline{\Phi}^i(T,\tau)\ \underline{B}\ d\tau\ \underline{y}^*(iT)$$

$$+ \sum_{i=o}^{\infty} \underline{\Phi}^i(T)\int_0^T \underline{\Phi}^i(T,\tau)\ \underline{\alpha}\ d\tau\ \underline{f}(\underline{x}^*(iT)). \qquad (4.25)$$

Clearly (4.25) is of the same form as (3.11), that is, a linear system with a perturbation. Thus the solution of (4.25) is obtained by casting it into the form of (3.28), the discrete version of (3.28) is

$$\left|\underline{\Gamma}(\underline{x}_2^*) - \underline{\Gamma}(\underline{x}_1^*)\right| = \left|\sum_{i=o}^{\infty} \underline{\Phi}^i(T)\ \underline{E}(T)\left[\frac{\partial f_j}{\partial x_k}\right](\underline{u}_2 - \underline{u}_1)\right|$$

$$\leq \left|\sum_{i=o}^{\infty} \underline{\Phi}^i(T)\right| \sup_T \left|\underline{E}(T)\left[\frac{\partial f_j}{\partial x_k}\right]\right| \|\underline{u}_2 - \underline{u}_1\| \qquad (4.26)$$

where $\underline{E} = \int_0^T \underline{\Phi}(T,\tau)\ \underline{\alpha}\ dT$.

The subscript attached to the vector $\underline{x}^*$ indicates the degree of approximation, i.e. x_1^* is the 1^{st} approximation to the final solution $\underline{x}^*$.

Let

$$\underline{P} = \left| \sum_{i-o}^{\infty} \underline{\Phi}^{i}(T) \right| \tag{4.27}$$

the upper limit of the summation in (4.27) has been extended to infinity to cover the general case. However, the resulting summation is a geometric series, and summing over an infinite number of terms is easy to handle providing the ratio between successive terms is less than one. Thus (4.27) may be written as

$$\underline{P} \quad \underline{<} \sum_{i=o}^{\infty} \left| \underline{\Phi}^{i}(T) \right| \tag{4.28}$$

and a single element of the $\underline{P}$ matrix may be written as

$$P_{kj} \underline{<} \sum_{i=o}^{\infty} \left| \Phi_{kj}^{i}(T) \right| .$$

At the same time make the following substitution, that is

$$\underline{z}(T) = \sup_{T} \left| \underline{E}(T) \left[\frac{\partial f_j}{\partial x_k} \right] \right|$$

$$\underline{<} \sup_{T} \left| \underline{E}(T) \right| \left| \left[\frac{\partial f_j}{\partial x_k} \right] \right| ; \tag{4.29}$$

Therefore,

$$z_{jm}(T) \leq \sup_{T} \left| Ejk \right| \left| \frac{\partial f_j}{\partial x_k} \right|, \tag{4.30}$$

where $1 \leq m \leq k$

thus

$$\sup_{k} \left| \sum_{j=1}^{n} P_{kj} Z_{jm}(T) \right| \| \underline{u}_2 - \underline{u}_1 \|$$

$$\leq \sum_{i=o}^{\infty} \left| \underline{\Phi}^{i}(T) \right| \sup_{T} \left| \underline{E}(T) \right| \left| \left[\frac{\partial f_j}{\partial x_k} \right] \right| \| \underline{u}_2 - \underline{u}_1 \| , \tag{4.31}$$

and by the definition of the norm,

$$\sup_{k} \left| \sum_{j=1}^{n} P_{kj} Z_{jm}(T) \right| = \left\| \underline{P}\ \underline{Z}(T) \right\|. \tag{4.32}$$

We now have sufficient information to determine the fixed point and the contraction condition.

4.2 A Convergence Criterion for Discrete Systems

The equations which represent a continuous non-linear system and its solution are written as (4.1) and (4.2) respectively. Equation (4.2) is cast into its discrete form which is written in its final form as (4.25). It is then solved by successive approximation methods as outlined in the material preceeding (4.26). A convergence criterion for the series which results from the use of (4.26) has been developed by Rao and Christensen, [38], and is reproduced here in part.

The sequence developed converges to a unique fixed point $\underline{x}^*(\cdot)$ in a sphere $\underline{S}$ of radius U centred at $\underline{x}^*_o(\ )$, if the following conditions are satisfied:

(a) contraction condition, (2.117),

$$\|\underline{\Gamma}(\underline{x}_2^*) - \underline{\Gamma}(\underline{x}_1^*)\| \leq K \|\underline{x}_2^* - \underline{x}_1^*\| \ \forall \ \underline{x}_1^*, \underline{x}_2^* \ \varepsilon \ \underline{S}.$$

(b) fixed point condition, (2.118),

$$\|\underline{\Gamma}(\underline{x}_o^*) - \underline{x}_o^*\| < (1 - K)\ U$$

where $o<K<1$.

The fact that a Volterra series can represent a discrete data system in the region in which the contraction mapping principle shows that B.I.B.O. stability exists can be justified on the same basis as has been done for the continuous system, namely section 3.3.

The uniqueness of the Volterra series follows directly from the contraction mapping principle and analyticity of the system follows from the fact that the series obtained can be cast into the form

$$\underline{\overline{X}}^* \leq \sum_{n=o}^{\infty} \|\underline{h}_n\| \, \underline{\overline{Y}}^n$$

where $\|\underline{h}_n\|$ for the discrete case is defined as

$$\|\underline{h}_n\| = \sum_{k_1=o}^{\infty} \text{----} \sum_{k_n=o}^{\infty} \left|\underline{h}_n(k_u\text{------}k_n)\right|.$$

4.3 The Solution of A Non-linear Type "1" System

The theoretical work required for the determination of the stability bound of a type "1" system is detailed in section 4.1. A practical demonstration of the method will be offered in this section. The non-linear system represented by equation (3.132) will be used again. This will promote, among other things, a convenient comparison of system stability between the continuous version and the sampled data system with data hold, keeping in mind that both represent the same non-linear differential equation.

The bound on the output vector $\underline{x}^*$ will be determined via the fixed point condition (3.91) and the contraction condition (3.102); the actual bound is obtained by taking the supremum of the two results.

Equation (3.132) is rewritten here for convenience

$$\ddot{x} + \dot{x} + x - \alpha x^3 = y(t),$$

this equation is now rewritten in matrix vector form, the same as (3.133) that is,

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + |\alpha| \begin{bmatrix} 0 \\ 1 \end{bmatrix} x_1^3 + \begin{bmatrix} 0 \\ 1 \end{bmatrix} y(t),$$

which is of the same form as (4.1) i.e.

$$\dot{\underline{x}}(t) = \underline{A}\ \underline{x}(t) + \underline{B}\ \underline{y}(t) + \underline{\alpha}\ \underline{f}(\underline{x}(t)).$$

The material developed in sections 4.1 and 3.4 will be utilized here to obtain the solution of the Type "1" system version of equation (3.133). The solution of this equation, when considered as representing a sampled data system, is given in general terms by (4.25). Further it is noted that the $\underline{P}$ matrix is one of the principles used in the determination of the bound on $\underline{x}^*$ and if (3.42) and (4.27) are compared it is seen that they differ, among other things, in the argument of the state transition matrix. Therefore, if the argument (t-s) used in equations (3.136) - (3.139) is replaced by (T), where T is the sample time, then these equations may be used again in the determination of the $\underline{P}$ matrix.

Thus if we allow T to equal 1 second, then from (3.136)

$$\Phi_{11}(T) = e^{-\cdot 5T}(\text{Cos } \cdot 866T + \frac{1}{\sqrt{3}} \text{ Sin} \cdot 866T) = \cdot 658, \tag{4.33}$$

and from (3.137)

$$\begin{aligned} \Phi_{12}(T) = - \Phi_{21}(T) &= e^{-\cdot 5T} \text{Sin} \cdot 866T \\ &= \cdot 461, \end{aligned} \tag{4.34}$$

finally from (3.139)

$$\begin{aligned} \Phi_{22}(T) &= e^{-\cdot 5T}(\text{Cos} \cdot 866 - \frac{1}{\sqrt{3}} \text{ Sin} \cdot 866T) \\ &= \cdot 125 \ . \end{aligned} \tag{4.35}$$

The summation of (4.28) is over an infinite number of terms and it is noted from (4.33), (4.34) and (4.35) that each of the elements of

the $\underline{P}$ matrix is less than one. Therefore, because $\Phi_{kj}(T) < 1$ in (4.28) then

$$P_{kj} \leq \frac{1}{1 - \Phi_{kj}}$$

Now substitute the values obtained in (4.33), (4.34), and (4.35) into (4.36), in succession. Thus values for the elements of the $\underline{P}$ matrix are now determined and found to be

$$\begin{aligned} P_{11} &\leq 2{\cdot}92 \\ P_{12} &\leq 1{\cdot}86 \\ P_{21} &\leq 0{\cdot}683 \\ P_{22} &\leq 1{\cdot}14 \end{aligned} \qquad (4.37)$$

As a result of this, the $\underline{P}$ matrix is written as

$$\underline{P} \leq \begin{bmatrix} 2{\cdot}92 & 1{\cdot}86 \\ 0{\cdot}683 & 1{\cdot}14 \end{bmatrix} . \qquad (4.38)$$

The $\underline{Z}$ matrix will now be determined from (4.29). Elements of the $\underline{E}$ matrix will be determined first, where

$$\begin{aligned} \left|\underline{E}(T)\right| &= \left|\underline{\alpha} \int_0^T \underline{\Phi}(T,\tau)\, d\tau\right| \\ &\leq \left|\underline{\alpha}\right| \int_0^T \left|\underline{\Phi}(T,\tau)\right| d\tau . \end{aligned} \qquad (4.39)$$

Therefore,

$$|E_{ij}(T)| \leq |\alpha| \int_0^T |\Phi_{ij}(T,\tau)| d\tau \qquad (4.40)$$

thus

$$|E_{11}(T)| \leq |\alpha| \int_0^T |\Phi_{11}(T,\tau)| d\tau. \qquad (4.41)$$

Which is of the same form as (3.140). Therefore, the material developed in Chapter 3 will be utilized again, starting at (3.141) and the ensuing development through to (3.146). Therefore

$$E_{11}(T) \leq |\alpha| \; Re^{-a\theta/b} \int_{\theta/b}^{T+\theta/b} |e^{az} \text{Sinbz}| dz, \qquad (4.42)$$

the values of a, θ, R and b, are as given in (3.142) which is

$R = 2/\sqrt{3}$

$b = \sqrt{3}/2$

$\theta = 1{\cdot}05$

$a = -{\cdot}5,$

thus,

$$E_{11}(T) \leq |\alpha| \; 1{\cdot}15e^{{\cdot}6} \int_{1{\cdot}21}^{2{\cdot}21} |e^{-{\cdot}5z} \text{Sin}{\cdot}866z| dz. \qquad (4.43)$$

Because the upper limit is less than $\pi/b = 3{\cdot}63$ the absolute value signs under the integral may be removed; then

$$E_{11}(T) \leq |\underline{\alpha}|\, 2{\cdot}10 \int_{1\cdot 21}^{2\cdot 21} e^{-\cdot 5z} \mathrm{Sin}\cdot 866z\,dz$$

$$\leq \cdot 872\ |\underline{\alpha}| \ . \tag{4.44}$$

Similarly with the remaining terms of the $\underline{E}$ matrix namely

$$E_{21}(T) = E_{12}(T) \leq \cdot 296\ |\underline{\alpha}| \tag{4.45}$$

$$E_{22}(T) \leq \cdot 530 |\underline{\alpha}| . \tag{4.46}$$

The values of the elements of the $\underline{E}$ matrix have now be determined, thus

$$\underline{E}(T) \leq \begin{bmatrix} \cdot 872 & \cdot 296 \\ \cdot 296 & \cdot 530 \end{bmatrix} |\underline{\alpha}| \tag{4.47}$$

Assuming that $|\underline{\alpha}| \equiv$ unit matrix, then the $\underline{z}$ matrix can be written as

$$\underline{z}(T) \leq \sup_{T} \begin{bmatrix} \cdot 872 & \cdot 296 \\ \cdot 296 & \cdot 530 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 3x^{*2} & 0 \end{bmatrix}$$

$$\leq \sup_{T} \begin{bmatrix} \cdot 888\underline{x}^{*2} & 0 \\ 1\cdot 59\underline{x}^{*2} & 0 \end{bmatrix} . \tag{4.48}$$

Now substitute (4.38) and (4.48) into (4.32), and write

$$\|\underline{P}\ \underline{z}(T)\| \leq \sup \left| \begin{bmatrix} 2\cdot 92 & 1\cdot 86 \\ 0\cdot 68 & 1\cdot 14 \end{bmatrix} \begin{bmatrix} \cdot 888\underline{x}^{*2} & 0 \\ 1\cdot 59\underline{x}^{*2} & 0 \end{bmatrix} \right|$$

$$\leq \sup \left| \begin{bmatrix} 5\cdot 56\underline{x}^{*2} \\ 2\cdot 42\underline{x}^{*2} \end{bmatrix} \right|$$

$$\leq \ 5\cdot 56\underline{x}^{*2} \ . \tag{4.49}$$

Thus

$$\|\underline{P}\ \underline{z}(T)\| \leq 5\cdot 56\underline{\bar{X}}^{*2} \ , \tag{4.50}$$

where

$$\underline{\bar{X}}^{*} = \sup_{i} \sup_{T} \left| x_i^*(T) \right| \ . \tag{4.51}$$

By comparing (4.50) with (3.55) it is clear that

$$5\cdot 56\underline{\bar{X}}^{*2} = K < 1, \tag{4.52}$$

the equation (4.52) has resulted from the contraction condition. The existence of a fixed point condition can now be determined from (3.102)

$$\sup_{T} \|\underline{P}\ \underline{L}(T)\| < (1-K)U. \tag{4.53}$$

The $\underline{P}$ matrix has already been determined (see inequality (4.38), and

$$\underline{L}(T) = \sup_{T} \left| \underline{E}(T)\underline{x}^{*3} \right| \ . \tag{4.54}$$

Substitute (4.47) into (4.54) and write

$$\underline{L}(T) \underline{<} \sup_{T} \left| \begin{bmatrix} \cdot 872 & \cdot 296 \\ \cdot 296 & \cdot 530 \end{bmatrix} \begin{bmatrix} 0 \\ \underline{x}^{*3} \end{bmatrix} \right|$$

$$\underline{<} \sup_{T} \begin{bmatrix} \cdot 296 & \underline{x}^{*3} \\ \cdot 530 & \underline{x}^{*3} \end{bmatrix} . \qquad (4.55)$$

Substituting (4.38) and (4.55) into (4.53) enables the L.H.S. of (4.53) to be evaluated, and consequently the R.H.S.; thus,

$$\|\underline{P}\ \underline{L}(T)\| \underline{<} \sup \left| \begin{bmatrix} 2\cdot 92 & 1\cdot 86 \\ 0\cdot 68 & 1\cdot 14 \end{bmatrix} \begin{bmatrix} \cdot 296\underline{x}^{*3} \\ \cdot 530\underline{x}^{*3} \end{bmatrix} \right|$$

$$\underline{<} \sup \left| \begin{matrix} 1\cdot 85\underline{x}^{*3} \\ 1\cdot 60\underline{x}^{*3} \end{matrix} \right|$$

$$\underline{<} \quad 1\cdot 85\underline{x}^{*3} \quad . \qquad (4.56)$$

By utilizing equation (4.51), inequality (4.56) may be rewritten as

$$\|\underline{P}\ \underline{L}(T)\| \underline{<} 1\cdot 85\overline{\underline{X}}^{*3} , \qquad (4.57)$$

and from (4.53)

$$1\cdot 85\overline{\underline{X}}^{*3} < (1-K)U . \qquad (4.58)$$

Inequalities (4.52) and (4.58) will now be used in the evaluation of $\underline{\bar{X}}^{*}$, from (4.52)

$$\underline{\bar{X}}^{*} \leq \left(\frac{K}{5{\cdot}56}\right)^{\frac{1}{2}} \tag{4.59}$$

and from (4.58)

$$\underline{\bar{X}}^{*} < \left(\frac{(1-K)U}{1{\cdot}855}\right)^{1/3}, \tag{4.60}$$

and if K, and U are assigned values that is let $K = {\cdot}6$ and $U = {\cdot}2$ then by the contraction condition, (4.59),

$$\underline{\bar{X}}^{*} \leq 0{\cdot}328 \tag{4.61}$$

and by the fixed point condition

$$\underline{\bar{X}}^{*} < 0{\cdot}348 \,. \tag{4.62}$$

If we now take the supremum of (4.61) and (4.62) it is easily seen that

$$\underline{\bar{X}}^{*} \leq {\cdot}328 \,. \tag{4.63}$$

Hence a bound on the output vector for a non-linear sampled data system, which is fitted with a sample hold device has been determined. Thus illustrating the applicability of the Volterra series to the solution of non-linear sampled data vector systems.

4.4 The Solution of a Non-linear Type "2" System

The determination of the stability region for a Type "2" system will be offered in this section. At the beginning of the section a general formula will be developed from which stability bounds can be determined. This will be followed by a practical demonstration of the method. The worked example will again use the same basic system equation as has been used in the earlier worked examples, namely equation (3.132), that is

$$\ddot{x} + \dot{x} + x - \alpha x^3 = y(t).$$

It has been shown previously that this equation may be written in vector matrix form as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + |\alpha| \begin{bmatrix} 0 \\ 1 \end{bmatrix} x_1^3 + \begin{bmatrix} 0 \\ 1 \end{bmatrix} y(t)$$

which was cited earlier as equation (3.133).

In section 4.3 it was noted that this equation was of the same form as (4.1), i.e.

$$\underline{\dot{x}}(t) = \underline{A}\ \underline{x}(t) + \underline{B}\ \underline{y}(t) + \underline{\alpha}\ \underline{f}(\underline{x}(t)).$$

This example is used again for a number of reasons, the most important of which is that it will provide us with a convenient means of comparing stability regions for various configurations of the same system.

A general formula for the stability region of a type "1" system has been obtained in section 4.1, and because of the basic similarities between a type "1" and type "2" system a large proporation of the material developed in that section will be utilized again here. The determination of a general formula for the stability bound in $\underline{x}^*$ for the type "2" system will therefore be initiated from equation (4.12). The material preceeding this equation is not reproduced in this section.

The data hold is not present in the type "2" system. Therefore, the input signal of the linear plant is a series of discrete finite pulses spaced "T" seconds apart. Thus the input signal to the linear plant at time "t" is

$$(\underline{B}\ \underline{y}^*(nT) + \underline{\alpha}\ \underline{f}(\underline{x}^*(nT)))\,\delta(t=nT) \tag{4.64}$$

where

$$\delta(t-nT) = \begin{cases} 1 \text{ for } t=nT \\ 0 \text{ for } t\neq nT. \end{cases} \tag{4.65}$$

Now expand (4.64) and rewrite as

$$\underline{B}\ y^*(nT)\ \ \delta(t-nT) + \underline{\alpha}\ \underline{f}(\underline{x}^*(nT))\ \ \delta(t-nT), \tag{4.66}$$

then by substituting (4.66) into (4.12) and applying equations (4.15), (4.16) and (4.8), where necessary, results in (4.12) being modified and we are able to rewrite it as

$$\underline{x}^*((n+1)T) = \underline{\Phi}(T)\ \underline{x}^*(nT) + \int_0^T \underline{\Phi}(T,\tau)\ \underline{\alpha}\ \underline{f}(\underline{x}^*(\tau+nT))\ \delta(\tau)d\tau$$

$$+ \int_0^T \underline{\Phi}(T,\tau)\ \underline{B}\ \underline{y}^*(\tau+nT)\ \delta(\tau)d\tau. \qquad (4.67)$$

Because of the delta function (4.65), equation (4.67) reduces to

$$\underline{x}^*((n+1)T) = \underline{\Phi}(T)\ \underline{x}^*(nT) + \Phi(T)\ \underline{\alpha}\ \underline{f}(x^*(nT))$$

$$+ \underline{\Phi}(T)\ \underline{B}\ \underline{y}^*(nT). \qquad (4.68)$$

Now modify equation (4.68) such that the initial conditions, i.e. $\underline{x}^*(0)$, are related to $\underline{x}^*((n+1)T)$ - see equation (4.24). Therefore,

$$\underline{x}^*(mT) = \underline{\Phi}^m(T)\ \underline{x}^*(0) + \sum_{i=o}^{m-1} \underline{\Phi}^{m-i}(T)\ \underline{B}\ \underline{y}^*(iT)$$

$$+ \sum_{i-o}^{m-1} \underline{\Phi}^{m-i}(T)\ \underline{\alpha}\ \underline{f}(\underline{x}^*(iT)). \qquad (4.69)$$

By extending the upper limit of the summatioms to infinity then (4.69) is rewritten as

$$\underline{x}^*(mT) = \underline{\Phi}^m(T)\ \underline{x}^*(0) + \sum_{i=o}^{\infty} \underline{\Phi}^{i}(T)\ \underline{B}\ \underline{y}^*(iT)$$

$$+ \sum_{i=o}^{\infty} \underline{\Phi}^{i}(T)\ \underline{\alpha}\ \underline{f}(\underline{x}^*(iT)), \qquad (4.70)$$

and by comparing (4.70) with (4.25) is is seen that they differ only in the following integral,

$$\int_0^T \underline{\Phi}(T,\tau)d\tau. \tag{4.71}$$

Clearly then (4.70) is a special case of equation (4.25). The method used to solve (4.25) will be utilized again except that the $\underline{E}$ matrix is replaced by the unit matrix. Thus equation (4.70) may now be written in its solution form

$$\left|\underline{\Gamma}(\underline{x}_2{}^*) - \underline{\Gamma}(\underline{x}_1{}^*)\right| = \left|\sum_{i=o}^{\infty} \underline{\Phi}^i(T)\ \underline{\alpha} \left[\frac{\partial f_k}{\partial x_j}\right] (\underline{u}_2 - \underline{u}_1)\right|$$

$$\leq \left|\sum_{i=o}^{\infty} \underline{\Phi}^i(T)\right| \left|\sup_T\right| \underline{\alpha} \left[\frac{\partial f_k}{\partial x_j}\right]\Bigg| \left|\left|\underline{u}_2 - \underline{u}_1\right|\right|. \tag{4.72}$$

Let

$$\underline{P} = \left|\sum_{i-o}^{\infty} \underline{\Phi}^i(T)\right|$$

which is identical to (4.27). Thus the elements of the $\underline{P}$ matrix take the same values as written in (4.38). The $\underline{z}$ matrix must now be determined, and if (4.72) is compared to (4.26) it is seen that they differ only in the $\underline{E}$ matrix, that is, replace the $\underline{E}$ matrix of (4.26) by the unit matrix and you obtain (4.72). Therefore, the $\underline{z}$ matrix may be

determined by using (4.29) and allowing $\underline{E} \equiv 1$.

Thus,

$$\underline{z}(T) \leq \sup_{T} \left| \left[\frac{\partial f_j}{\partial x_k} \right] \right|$$

which for the example in question can be rewritten as

$$\underline{z}(T) \leq \sup_{T} \begin{bmatrix} 0 & 0 \\ 3\underline{x}^{*2} & 0 \end{bmatrix} |\underline{\alpha}| \tag{4.73}$$

and if we allow $|\underline{\alpha}| \equiv \underline{1}$ then enough information exists to enable us to determine (4.32), which is rewritten here for convencience

$$\sup_{k} \left| \sum_{j=1} P_{kj} z_{jm}(T) \right| = \| \underline{P} \; \underline{z}(T) \| .$$

Therefore,

$$\| \underline{P} \; \underline{z}(T) \| = \sup \left| \begin{bmatrix} 2{\cdot}92 & 1{\cdot}85 \\ {\cdot}068 & 1{\cdot}14 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 3\underline{x}^{*2} & 0 \end{bmatrix} \right|$$

$$\leq \sup \left| \begin{matrix} 5{\cdot}59\underline{x}^{*2} \\ 3{\cdot}42\underline{x}^{*2} \end{matrix} \right|$$

$$\leq 5{\cdot}59\underline{x}^{*2} . \tag{4.74}$$

Thus

$$\| \underline{P} \; \underline{z}(T) \| \leq 5{\cdot}59\underline{\bar{X}}^{*2} , \tag{4.75}$$

and from (3.55) it is clear that

$$5{\cdot}59\,\underline{\overline{X}}^{*2} = K<1, \tag{4.76}$$

which is a statement of the contraction condition.

The fixed point condition is easily obtained from (3.99), that is

$$\sup_T \| \underline{P}\ \underline{L}(T) \| < (1-K)U\,, \tag{4.77}$$

The $\underline{P}$ matrix is already known and the $\underline{L}$ vector can be determined from (4.54) and (4.55) remembering that $\underline{E} \equiv \underline{1}$. Therefore,

$$\underline{L}(T) \leq \sup_T \left| \begin{bmatrix} 0 \\ \underline{x}^{*3} \end{bmatrix} \right| . \tag{4.78}$$

Thus the L.H.S. of (4.77) can be evaluated as

$$\begin{aligned} \| \underline{P}\ \underline{L}(T) \| &= \sup \left| \begin{bmatrix} 2{\cdot}92 & 1{\cdot}86 \\ 0{\cdot}68 & 1{\cdot}14 \end{bmatrix} \begin{bmatrix} 0 \\ \underline{x}^{*3} \end{bmatrix} \right| \\ &\leq \sup \left| \begin{bmatrix} 1{\cdot}86\,\underline{x}^{*3} \\ 1{\cdot}14\,\underline{x}^{*3} \end{bmatrix} \right| \\ &\leq 1{\cdot}86\,\underline{x}^{*3}\,, \end{aligned} \tag{4.79}$$

and from (4.77)

$$1{\cdot}86\,\underline{\overline{X}}^{*3} < (1-K)\,U\,. \tag{4.80}$$

Equation (4.76) can now be rewritten as

$$\underline{\overline{X}}^* \leq \left(\frac{K}{5 \cdot 59}\right)^{\frac{1}{2}} \quad . \tag{4.81}$$

Similarly from equation (4.80)

$$\underline{\overline{X}}^* < \left(\frac{(1-K)U}{1 \cdot 86}\right)^{1/3} , \tag{4.82}$$

if we now allow

$$K = \cdot 6$$

$$U = \cdot 2$$

then from the contraction condition (4.81)

$$\underline{\overline{X}}^* \leq \cdot 327 , \tag{4.83}$$

and from the fixed point condition (4.82)

$$\underline{\overline{X}}^* < \cdot 347 \tag{4.84}$$

and by taking the supremum of (4.83) and (4.84) it is clearly seen that

$$\underline{\overline{X}}^* \leq \cdot 327 . \tag{4.85}$$

Thus a bound on the output vector for a type "2" system has been determined.

Conclusions On Sections 4.3 and 4.4:

The bounds on $\underline{\overline{X}}^*$ are smaller than the corresponding bounds for a continuous system. This is to be expected since a sampled data system tends to make the system less stable than a continuous data system.

4.5 Inbetween Sampling of Type "1" and Type "2" Systems

A knowledge of the behaviour of a system between sampling instants is important to the systems engineer. Some systems which are designed to have a finite settling time at sample instants do not necessarily settle to an equilibrium condition between samples. That is, undershoots and overshoots may occur between samples and not be detected. One of the important advantages of the state variable method over the z transform method is that it can be easily modified. Thus the state and output of the system are describable between sample periods.

A portion of figure 4 is reproduced here to assist in illustrating the method utilized in the determination of the output stability region.

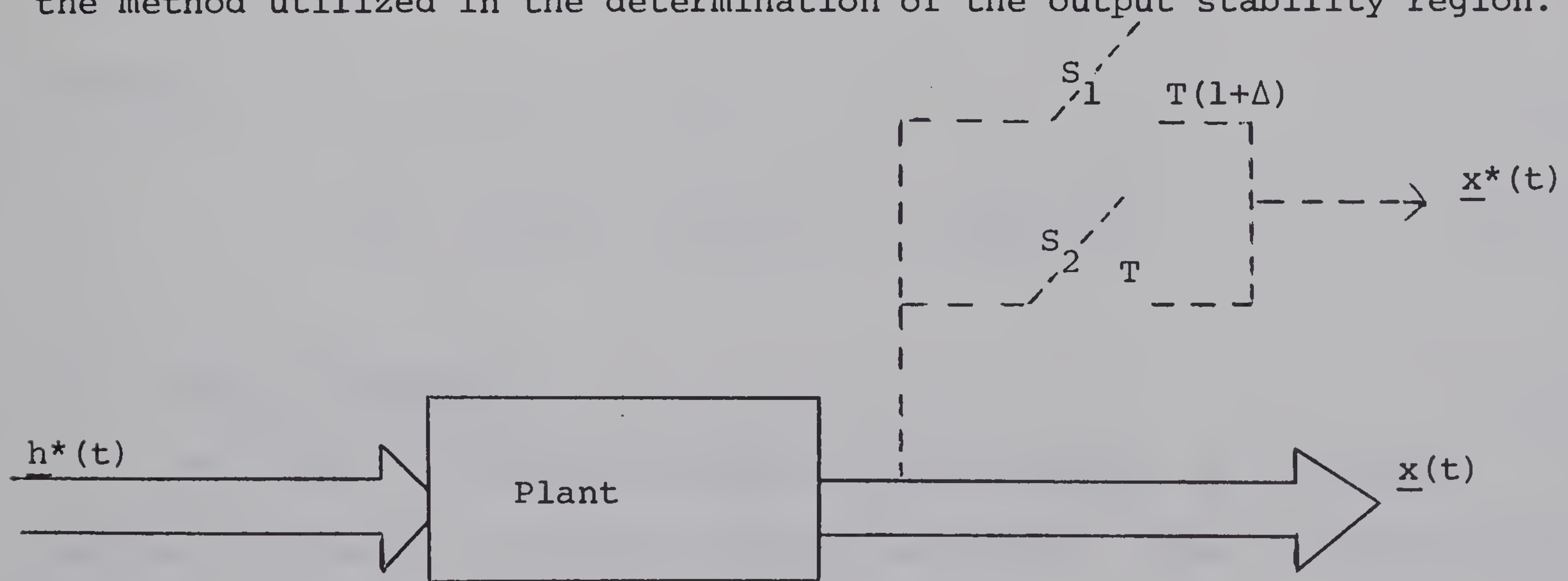

Figure 8: Model Used for the Determination of the Inter Sample Stability Region

The type "1" and type "2" systems previously utilized in sections 4.1, 4.3 and 4.4 will be used again, in this section, with system

equation (3.1). This repetitious use of (3.1) will illustrate clearly what is happening in the stability region of a data system as opposed to a continuous system.

When considering in between sampling the system is open loop as shown, in Fig. 6. Actually what is required by the systems designer is the bound on the continuous system output, $\underline{x}(t)$.

Consider then a particular sampling instant "n", thus

$$\dot{\underline{x}}(t) = \underline{A}\ \underline{x}(t) + \underline{h}^*(nT), \tag{4.86}$$

where

$$nT < t \leq (n+1)T. \tag{4.87}$$

But

$$\underline{h}^*(nT) = \underline{B}\ \underline{y}^*(nT) + \underline{\alpha}\ \underline{f}(\underline{x}^*(nT)), \tag{4.88}$$

therefore,

$$\dot{\underline{x}}(t) = \underline{A}\ \underline{x}(t) + \underline{B}\ \underline{y}^*(nT) + \underline{\alpha}\ \underline{f}(\underline{x}^*(nT)). \tag{4.89}$$

4.5.1 Type "1" System

The input to the plant, between sample periods, is a constant (section 4.1 equations (4.12) to (4.17) inclusive). Thus the solution of the system equation by standard state space methods is;

$$\underline{x}(t) = \underline{\Phi}(t,t_o)\underline{x}(t_o) + \underline{y}(t_o)\int_{t_o}^{t} \underline{\Phi}(t,s)\underline{B}ds$$

$$+ \underline{f}(\underline{x}(t_o))\int_{t_o}^{t} \underline{\Phi}(t,s)\ \underline{\alpha}\ ds. \tag{4.90}$$

Now let

$$t = (n+\Delta)T$$

$$t_o = nT \qquad (4.91)$$

where $o<\Delta\underline{<}1$,

and substitute (4.91) into (4.90) and write

$$\underline{x}^*((n+\Delta)T) = \underline{\Phi}(\Delta T)\underline{x}^*(nT) + \underline{y}^*(nT)\int_{nT}^{(n+\Delta)T} \underline{\Phi}((n+\Delta)T,s)\underline{B}ds$$

$$+ \underline{f}(\underline{x}^*(nT))\int_{nT}^{(n+\Delta)T} \underline{\Phi}((n+\Delta)T,s)\ \underline{\alpha}\ ds. \qquad (4.92)$$

If now we let

$$nT - s = -\tau, \qquad (4.93)$$

then

$$\underline{\Phi}((n+\Delta)T,s) = \underline{\Phi}(\ T,\tau); \qquad (4.94)$$

therefore,

$$\int_{nT}^{(n+\Delta)T} \underline{\Phi}((n+\Delta)T,s){-}{-}{-}{-}ds = \int_{0}^{\Delta T} \underline{\Phi}(\Delta T,\tau){-}{-}d\ . \qquad 4.95)$$

Thus equation (4.92) can now be rewritten as

$$\underline{x}^*((n+\Delta)T) = \underline{\Phi}(\Delta T)\underline{x}^*(nT) + \underline{y}^*(nT)\int_{0}^{\Delta T} \underline{\Phi}(\Delta T,\tau)\underline{B}d\tau$$

$$+ \underline{f}(\underline{x}^*(nT))\int_{0}^{\Delta T} \underline{\Phi}(\Delta T,\tau)\ \underline{\alpha}\ d\tau, \qquad (4.96)$$

and it is observed that (4.96) is analogous to (4.17) - as expected. Also it is now possible to determine the bound on $\underline{x}(t)$ - after the n^{th} sample - at anytime "t" in the interval

$$nT < t \leq (n+1)T ,$$

By comparing equations (4.96) and (3.11) it is easily seen that they too are analogous, therefore the solution of (4.96) will be obtained by casting it into the form of (3.28). That is

$$\left|\underline{\Gamma}(\underline{x}_2^*) - \underline{\Gamma}(\underline{x}_1^*)\right| = \left|\int_0^{\Delta T} \underline{\Phi}(\Delta T,\tau)\ \underline{\alpha} d\tau \left[\frac{\partial f_j}{\partial x_k}\right] (\underline{u}_2 - \underline{u}_1)\right|$$

$$\leq \sup_t \left|\int_0^{\Delta T} \underline{\Phi}(\Delta T,\tau)\ \underline{\alpha} d\tau \left[\frac{\partial f_j}{\partial x_k}\right]\right| \|\underline{u}_2 - \underline{u}_1\| . \qquad (4.97)$$

Now let

$$\underline{z}(t) = \sup_t \left|\int_0^{\Delta T} \underline{\Phi}(\Delta T,\tau)\ \underline{\alpha} d\tau\right| \left|\left[\frac{\partial f_j}{\partial x_k}\right]\right| , \qquad (4.98)$$

and

$$\underline{E}(\Delta T) = \int_0^{\Delta T} \underline{\Phi}(\Delta T,\tau)\ \underline{\alpha}\ d\tau ; \qquad (4.99)$$

therefore,

$$z_j(t) = \sup_t \sum_{k=1}^{N} \left|E_{jk}(\Delta T)\right| \left|\frac{\partial f_j}{\partial x_k}\right| . \qquad (4.100)$$

Thus,

$$\sup_{t}\left|\int_{0}^{\Delta T} \underline{\Phi}(\Delta T,\tau)\ \underline{\alpha}\ d\tau \left[\frac{\partial f_j}{\partial x_k}\right]\right| \| u_2 - u_1 \|$$

$$\leq \sup_{j}\left|z_j(t)\right| \| \underline{u}_2 - \underline{u}_1 \| , \tag{4.101}$$

and by the definition of norm

$$\sup_{j}\left|z_j(t)\right| = \| \underline{z}(t) \| . \tag{4.102}$$

It is to be remembered that the in between sample stability is being sought for a type "1" system version of equation (3.1), with sample period 1 second. Material that was previously developed for the determination of sample stability will be used again. The elements of the E matrix will be determined first. This willbe followed by the determination of the elements of the $\underline{z}$ vector. Formulas, inequalities, etc. appropriate to this, and previously developed, will be utilized again.

The elements of the $\underline{E}$ matrix will now be evaluated starting with $E_{11}(T)$, (4.42)

$$E_{11}(T) \leq | \underline{\alpha} | \ Re^{-a\theta/b}\int_{\theta/b}^{T+\theta/b} \left|e^{az}\text{Sinbz}\right| dz$$

and replacing T by ΔT in this inequality as in (4.99) results in (4.42) being written as

$$E_{11}(\Delta T) \leq | \ \alpha \ | \ Re^{-a\theta/b}\int_{\theta/b}^{\Delta T+\theta/b} \left|e^{az}\text{Sinbz}\right| dz. \tag{4.103}$$

Clearly (4.42) is a special case of (4.103), when $\Delta = 1$. By similar methods $E_{12}(\Delta T)$, $E_{21}(\Delta T)$ and $E_{22}(\Delta T)$ will be determined. The magnitudes of the elements of the $\underline{E}$ matrix are listed below for various values of Δ, where T = 1sec.

Δ	$\lvert E_{11}\rvert$	$\lvert E_{12}\rvert = \lvert E_{21}\rvert$	$\lvert E_{22}\rvert$	
·1	·099	·004	·094	
·2	·198	·016	·171	
·3	·295	·035	·254	
·4	·389	·060	·320	
·5	·480	·090	·376	
·6	·568	·125	·423	(4.104)
·7	·651	·163	·462	
·8	·723	·205	·492	
·9	·803	·250	·515	
1·0	·872	·294	·531	

thus, $\underline{\bar{x}}^*$ can now be determined.

By the contraction condition

$$\|\underline{a}(t)\| \leq \sup_t \sum_{k=1}^{2} \left| \begin{bmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 3x^{*2} & 0 \end{bmatrix} \right|$$

$$\leq \sup_t \left| \begin{matrix} 3E_{12}\underline{x}^{*2} \\ 3E_{22}\underline{x}^{*2} \end{matrix} \right| \tag{4.105}$$

$$\leq \sup \left| \begin{matrix} 3E_{12} \\ 3E_{22} \end{matrix} \right| \underline{\bar{X}}^{*2} \leq K < 1 \; .$$

By the fixed point condition

$$\|\underline{z}(t)\| \leq \sup_{t} \left| \begin{bmatrix} E_{11} & E_{12} \\ E_{21} & E_{22} \end{bmatrix} \begin{bmatrix} 0 \\ x^{*3} \end{bmatrix} \right| \tag{4.106}$$

$$\leq \sup \left| \begin{matrix} E_{12} \\ E_{22} \end{matrix} \right| \underline{\bar{X}}^{*3} < (1 - K)U.$$

For $K = \cdot 6$ and $U = \cdot 2$ a table of bounds on $\underline{\bar{X}}^{*}$ for various values of Δ are listed below,

Δ	Fixed Point	Contraction	$\underline{\bar{X}}^{*}$
·1	·944	1·45	·944
·2	·764	1·05	·764
·3	·680	·887	·680
·4	·630	·790	·630
·5	·597	·730	·597
·6	·574	·687	·574
·7	·557	·658	·557
·8	·545	·637	·545
·9	·537	·622	·537
1·0	·531	·613	·531 .

(4.107)

4.5.2 <u>Type "2" System</u>

The system is again considered as an open loop system, as in Figure 6, the input is a series of discrete pulses and the solution equation is

$$\underline{x}(t) = \underline{\Phi}(t,t_o)\underline{x}(t_o) + \int_{t_o}^{t} \underline{\Phi}(t,s)\ \underline{B}\ \underline{y}(s)\underline{\delta}(s-t_o)ds$$

$$+ \int_{t_o}^{t} \underline{\Phi}(t,s)\ \underline{\alpha}\ \underline{f}(\underline{x}(s))\underline{\delta}(s-t_o)ds. \qquad (4.108)$$

Because of the delta function equation (4.108) reduces to

$$\underline{x}(t) = \underline{\Phi}(t,t_o)\underline{x}(t_o) + \underline{\Phi}(t,t_o)\underline{B}\ \underline{y}(t_o)$$

$$+ \underline{\Phi}(t,t_o)\ \underline{\alpha}\ \underline{f}(\underline{x}(t_o)). \qquad (4.109)$$

If we now let

$$\begin{aligned} t &= (n + \Delta)T \\ t_o &= nT \end{aligned} \qquad (4.110)$$

and substitute (4.110) into (4.109) then,

$$\underline{x}^*((n + \Delta)T) = \underline{\Phi}(\Delta T)\underline{x}^*(nT) + \underline{\Phi}(\Delta T)\ \underline{B}\ \underline{y}^*(nT)$$

$$+ \underline{\Phi}(\Delta T)\ \underline{\alpha}\ \underline{f}(\underline{x}^*(nT)). \qquad (4.111)$$

By methods previously used the solution of (4.111) is determined as follows,

$$\left|\underline{\Gamma}(\underline{x}_2{}^*) - \underline{\Gamma}(\underline{x}_1{}^*)\right| = \left|\underline{\Phi}(\Delta T)\ \underline{\alpha}\left[\frac{\partial f_j}{\partial x_k}\right](\underline{u}_2 - \underline{u}_1)\right|$$

$$\leq \sup_t \left|\underline{\Phi}(\Delta T)\ \underline{\alpha}\left[\frac{\partial f_j}{\partial x_k}\right]\right| \|\underline{u}_2 - \underline{u}_1\| . \tag{4.112}$$

Now make the following substitutions into (4.112), that is let the matrix

$$\underline{z}(t) = \sup_t |\underline{\alpha}| \left|\left[\frac{\partial f_j}{\partial x_k}\right]\right| \tag{4.113}$$

and the matrix

$$\underline{P} = \left|\ \underline{\Phi}(\Delta T)\right| , \tag{4.114}$$

where

$$z_{jm}(t) \leq \sup_t \left|\alpha_j{}^k\right| \left|\left[\frac{\partial f_j}{\partial x_k}\right]\right| \tag{4.115}$$

and

$$P_{kj} \leq \left|\Phi_{kj}(\Delta T)\right| . \tag{4.116}$$

Thus

$$\sup_{t}\left|\underline{\Phi}(\Delta T)\ \underline{\alpha}\ \left[\frac{\partial f_j}{\partial x_k}\right]\right| \, ||\underline{u}_2 - \underline{u}_1||$$

$$\leq \sup_{k}\left|\sum_{j=o}^{N} P_{kj} z_{jm}(t)\right| \, ||\underline{u}_2 - \underline{u}_1|| \ , \tag{4.117}$$

and by the definition of a norm

$$||\underline{P}\ \underline{z}(t)|| = \sup_{k}\left|\sum_{j=o}^{N} P_{kj} z_{jm}(t)\right| . \tag{4.118}$$

The bound on $\underline{x}^*((n+\Delta)T)$ will now be determined, from equation (3.136)

$$\Phi_{11}(\ T) = e^{-\cdot 5\Delta T}(\mathrm{Cos}\cdot 866\Delta T + \frac{1}{\sqrt{3}}\ \mathrm{Sin}\cdot 866\Delta T), \tag{4.119}$$

similarly with $\Phi_{12}(\Delta T)$, $\Phi_{21}(\Delta T)$ and $\Phi_{22}(\Delta T)$.

The magnitudes of the elements of the $\underline{P}$ matrix are listed below for various values of Δ, where $T = 1\mathrm{sec}$,

Δ	$\lvert P_{11}\rvert$	$\lvert P_{12}\rvert=\lvert P_{21}\rvert$	$\lvert P_{22}\rvert$
·1	·994	·086	·900
·2	·980	·172	·800
·3	·958	·257	·700
·4	·930	·339	·610
·5	·893	·419	·520
·6	·853	·496	·432
·7	·808	·569	·349
·8	·760	·638	·270
·9	·710	·702	·197
1·0	·657	·761	·128 .

(4.120)

The elements of the $\underline{z}$ matrix are obtainable from (4.73), and (4.78). The elements of the $\underline{L}$ vector are obtainable from (4.78).

By the contraction condition

$$\|\underline{P}\ \underline{z}(t)\| \leq \sup_t \left| \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 3\underline{x}^{*2} & 0 \end{bmatrix} \right|$$

$$\leq \sup_t \left| \begin{matrix} 3P_{12}\underline{x}^{*2} \\ 3P_{22}\underline{x}^{*2} \end{matrix} \right| \tag{4.121}$$

$$\leq \sup \left| \begin{matrix} 3P_{12} \\ 3P_{22} \end{matrix} \right| \underline{\bar{X}}^2 < K < 1.$$

By the fixed point condition

$$\|\underline{P}\ \underline{L}(t)\| \leq \sup_{t} \left| \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} 0 \\ \underline{x}^{*3} \end{bmatrix} \right|$$

$$\leq \sup \left| \begin{matrix} P_{12} \\ P_{22} \end{matrix} \right| \underline{\bar{X}}^{*3} < (1 - K)u. \qquad (4.122)$$

The table of bounds on $\underline{\bar{X}}^{*}$ for various values of Δ, with $K = \cdot 6$ and $u = \cdot 2$, is listed below

Δ	Fixed Point	Contraction	$\underline{\bar{X}}^{*}$	
·1	·471	·446	·446	
·2	·500	·463	·463	
·3	·532	·484	·484	
·4	·572	·508	·508	
·5	·620	·535	·535	
·6	·680	·570	·570	(4.123)
·7	·700	·584	·584	
·8	·680	·571	·571	
·9	·668	·563	·563	
1·0	·658	·557	·557.	

In between sample stability has shown that the system is stable between sample periods, as exhibited by the intersample bounds $\underline{\bar{X}}^{*}$ of (4.123).

The use of the Volterra series as a tool in the solution of N.L. sampled data systems has been demonstrated. The technique has shown

that providing the parent equation can be written in the form of (3.1) and is analytic then a Volterra series solution exists, if the contraction and fixed point conditions are satisfied.

4.6 The Determination of the Bound on the Input Vector

The bound on the input vector $\underline{y}^*(t)$ will now be determined via the method presented in the latter part of section 3.4.2, commencing at (3.225) namely,

$$\underline{x}^* = \underline{\ell} + \underline{u}^* \; . \tag{4.124}$$

By taking the norm of (4.124) it can now be written as

$$\underline{\overline{X}^*} \;\underline{<}\; \mathcal{L} + U \tag{4.125}$$

and from the material in sections 3.1 and 3.2 it is clearly seen that

$$\mathcal{L} \;\underline{<}\; \sup \left| \underline{P}\;\underline{Y} \right| \; . \tag{4.126}$$

As an illustration of the method one example will be cited, namely the system considered in section 4.1.

Thus $\mathcal{L}$ will be determined from the $\underline{P}$ matrix (4.38) and the $\underline{Y}$ vector. Therefore, $\mathcal{L}$ can be written as

$$\mathcal{L} \;\underline{<}\; \sup \left| \begin{bmatrix} 2{\cdot}92 & 1{\cdot}86 \\ 0{\cdot}683 & 1{\cdot}14 \end{bmatrix} \begin{bmatrix} 0 \\ y \end{bmatrix} \right| \tag{4.127}$$

$$\underline{<}\; 1{\cdot}86\underline{\overline{Y}} \tag{4.128}$$

where

$$\underline{\bar{Y}} \leq \sup_{i} \sup_{t} |y_i(t)| \ , \tag{4.129}$$

and by examination of (4.128) it is clearly seen that it may be rewritten as

$$\underline{\rho} \leq H\underline{\bar{Y}} \tag{4.130}$$

where $$H = 1{\cdot}86 \ , \tag{4.131}$$

then (4.125) can be rewritten as

$$\underline{\bar{X}}^* \leq H\underline{\bar{Y}} + U \ . \tag{4.132}$$

Substituting (4.132) into the contraction condition (4.59) results in

$$H\underline{\bar{Y}} + U \leq \left(\frac{K}{3|\alpha|H}\right)^{\frac{1}{2}} \tag{4.133}$$

and for substituting the values assumed for K, α, and U into (4.133), that is $K = {\cdot}6$, $U = {\cdot}2$ and $\alpha = 1$, then $\underline{\bar{Y}}$ is found to be

$$\underline{\bar{Y}} \leq \left(\left(\frac{{\cdot}6}{3\mathrm{x}1\mathrm{x}1{\cdot}86}\right)^{\frac{1}{2}} - {\cdot}2\right) \frac{1}{1{\cdot}86}$$

$$\underline{\bar{Y}} \leq {\cdot}068 \ . \tag{4.134}$$

Substituting (4.132) into the fixed point condition (4.60) results in

$$H\underline{\overline{Y}} + \mathbf{U} < \left(\frac{(1 - K)\mathbf{U}}{H|\alpha|}\right)^{1/3} \tag{4.135}$$

from which $\underline{\overline{Y}}$ is found to be

$$\underline{\overline{Y}} < .080 \quad . \tag{4.136}$$

By taking the supremum of (4.134) and (4.136) it is seen that

$$\underline{\overline{Y}} \leq .068 \quad . \tag{4.137}$$

Hence a bound on the input vector has been determined.

Bounds on the input vector for the type "2" system can be determined in a similar manner.

CHAPTER V

CONCLUSION

The thesis has utilized the convergence properties of the Volterra series via the Banach contraction mapping principle to prove that bounded input - bounded output stability does exist for a certain class of multivariate non-linear systems. To ensure that the gradient of the non-linearity does not exceed the design figure, the power of the Frechet derivative was enlisted.

The class of systems applicable to this thesis are those whose characterizing differential equations satisfy the following conditions:

(1) the differential equations can be cast into vector matrix form i.e.

$$\dot{\underline{x}}(t) = \underline{Ax}(t) + \underline{By}(t) + \underline{\alpha}\ \underline{f}(\underline{x}(t)).$$

(2) the differential equations are analytic.

It should be noted that although the material here is restricted to a certain class of systems, the class is extremely large and contains most of the practical systems.

Bounded input - bounded output stability is investigated in this thesis for one particular system. The equation for this system was considered as representing:

(a) a continuous system,

(b) a sampled data system with and without zero order hold.

Inter sample stability is considered for system (b).

It has been shown in section 3.3 that a Volterra series can be obtained by an iterative technique from a given system's non-linear vector matrix d.e. A convergence and a uniqueness criterion, for the

Volterra series, has been developed by Christensen [10] and [33], respectively. These criteria are based on the contraction mapping and fixed point principles and upon the definition of an analytic system expounded by Brilliant [7].

Based upon the methods developed by Christensen, the idea behind the thesis was to find, erect and demonstrate a simple procedure whereby once the vector matrix form of the non-linear d.e. was known, it would be comparatively easy to determine the region in which B.I.B.O. stability held, and that a Volterra series did exist. As previously stated two basic tests were developed by Christensen.

(a) the convergence test [10]

(b) the uniqueness test [33].

Test (a) determines if the representative Volterra series is convergent if so, how large is the region of convergence. The method of this test displayed in sections 3.4.1 and 3.4.2, from (3.184) to (3.192) and (3.212) to (3.216). Also test (b) determines if the representative Volterra series is unique, if so how large is the region of uniqueness. The method of this test is displayed in sections 3.4.1 and 3.4.2, from (3.184) to (3.193) and (3.212) to (3.221).

The region over which both tests can be made to apply simultaneously is determined by taking the supremum of the regions determined in test (a) and test (b). Once the "size" of the region is known, bounds may be applied to the input and output vectors.

Sufficient conditions only were established for the determination of the region of stability.

Actually how close are the calculated bounds to the actual solution? The accompanying curves show the response of the continuous system to

various step inputs. The equation was solved by the Runge Kutta method and plotted with the aid of a Honeywell 316 computer. From the curve it is clearly seen that the system is unstable if

$$|\underline{y}(t)| \leq \cdot 470. \tag{5.1}$$

For

$$|\underline{y}(t)| = \cdot 450 \tag{5.2}$$

the output

$$|\underline{x}(t)| \leq \cdot 611. \tag{5.3}$$

It must be conceded that the results obtained in Chapters 3 and 4 are conservative. Principally they are based on:

(1) the assumption that the non-linear feed back network is exerting maximum effort to make the system go unstable.

(2) generally it is only possible - as has been demonstrated in the thesis - to compute the bounds on the norm of the input/output vector.

(3) the sup norm has been used throughout, that is

$$\underline{\overline{X}}^{*} = \|\underline{x}(t)\| = \sup\left|x_i(t)\right|, \tag{5.4}$$

meaning that the "size" of the vector is represented by its largest component.

The bounds on the components of the continuous system are

$$\underline{\overline{X}}^{*} \leq \cdot 405 \tag{5.5}$$

$$\underline{\overline{Y}} \leq \cdot 173 \tag{5.6}$$

from the curve for

$$|\underline{y}(t)| = \cdot 173 \tag{5.7}$$

it is seen that

$$|\underline{x}(t)| \leq \cdot 188 \quad . \tag{5.8}$$

Clearly by comparing (5.2) with (5.7) and (5.3) with (5.8) it is seen that the results are conservative. However, in spite of this short-coming the method developed here is realistic in that not only is stability determined but also the "size" of the stability region is found.

Work is required to establish a method for predicting the advent of limit cycles once bounded input - bounded output stability has been determined. Also, a criterion needs to be established which will enable the systems designer to determine controllability and observability of the multivariable non-linear system.

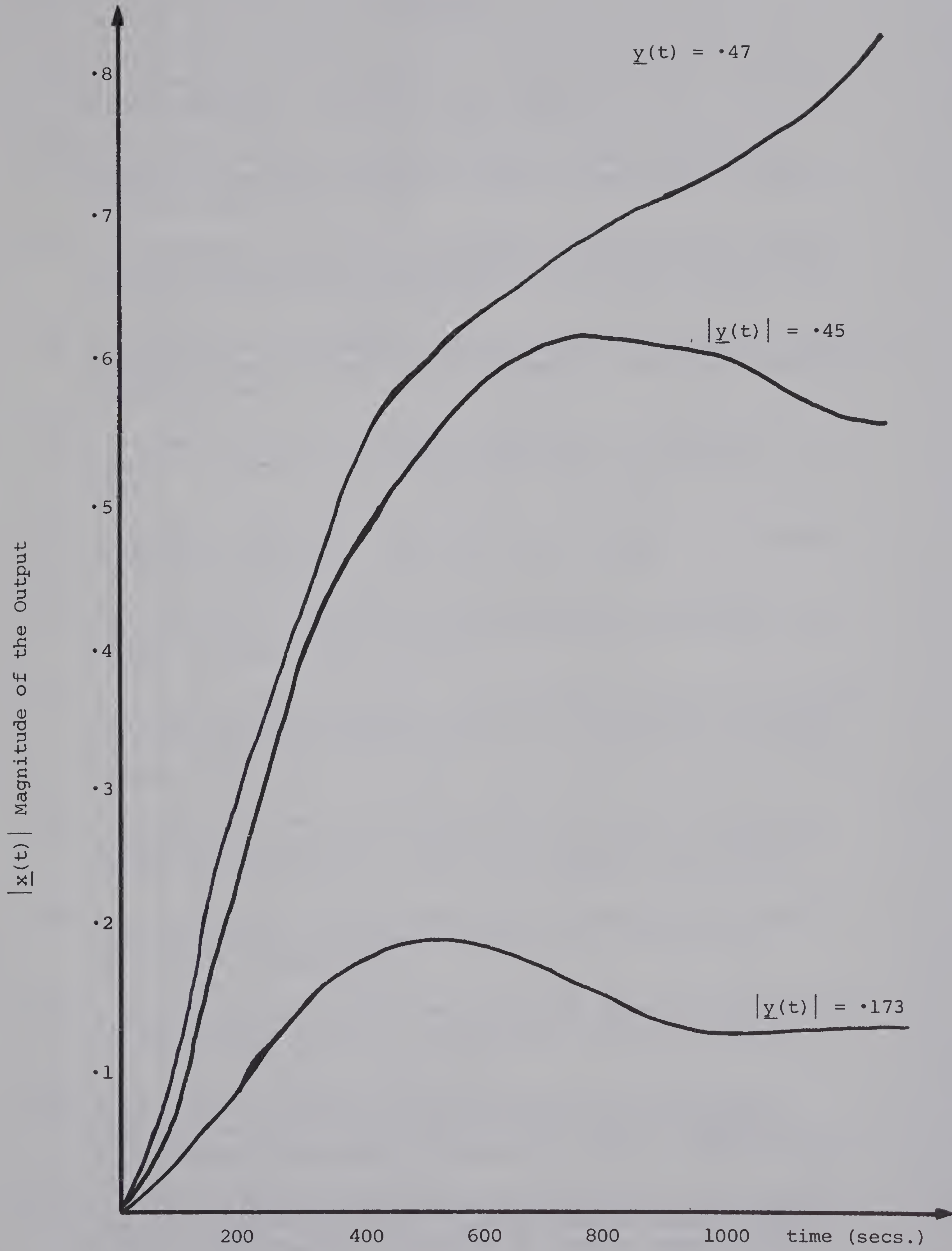

Figure 9: Response Curve for $\ddot{x} + \dot{x} + x - \alpha x^3 = y(t)$

REFERENCES

[1] J.L. Bower and P.M. Schultheiss, Introduction to the design of servo mechanisms. New York: Wiley, 1958.

[2] R.E. Kalman and J.E. Bertram, "Control system analysis and design via the second method of Liapunov, ASME Trans. J. Basic Eng. 82D, pp. 371-393, 1960.

[3] H. Antosiewicz, "A survey of Liapunov's second method", Contributions to non-linear oscillations IV, Ann. Math. Studies, Vol. 41, pp. 147-166. Princeton 1958.

[4] D.R. Ingwerson, "A modified Liapunov method for nonlinear stability problems", Doctoral Dissertation, Stanford University, Stanford, California 1960.

[5] L.A. Zadeh, A general theory of linear signal transmission systems. Journal of the Franklin Institute. pp. 293-312, April 1952.

[6] V. Volterra, Theory of functionals and of integral and integro-differential equations. New York: Dover, 1959.

[7] M.B. Brilliant, "Theory of analysis of nonlinear systems", Massachusetts Institute of Technology Electronics Research Lab., Cambridge, Mass., Rep. 345, 1959.

[8] R.H. Flake, "Volterra series representation of nonlinear systems", Transactions of the American Institute of Electrical Engineers, Part II; Applications and industry, Vol. 81, No. 6, pp. 330-335, January 1963.

[9] J.F. Barrett, "The use of Volterra series to find a region of stability of a nonlinear differential equation", International Journal of Control, Vol. 1, No. 3, pp. 209-216, March 1965.

[10] G.S. Christensen, "On the convergence of the Volterra series", IEEE Transactions on Automatic control, Vol. AC-13, No. 6, pp. 736-737, December 1968.

[11] G.W. Trott and G.S. Christensen, "A larger region of contraction for the Volterra series". International Journal of Control, Vol. 14, No. 2, pp. 337-384, 1971.

[12] A.M.H. Rashed, "Studies of nonlinear discrete data systems with non-zero initial conditions", M.Sc. thesis, Department of Electrical Engineering, University of Alberta, Edmonton 1970.

[13] R.S. Rao, "Stability studies of nonlinear discrete data systems through functional analysis", Doctoral thesis, Department of Electrical Engineering, University of Alberta, Edmonton 1969.

[14] S.C. Chu and J.B. Biaz, "On 'In the large' application of the contraction principle", Atti Della Accademia Delle Scienze Di Torino, Vol. 99, pp. 351-363, 1964-1965.

[15] P.L. Falb and J.L. DeJong, Some successive approximation methods in control and oscillation theory. New York: Academic 1969.

[16] W.A. Porter, Modern foundations of systems engineering. New York: Macmillan, 1966.

[17] M. Athans and P.L. Falb, Optimal control: An introduction to the theory and its applications. New York: McGraw-Hill, 1966.

[18] M.Z. Nashed, "Role of differentials in non-linear functional analysis". Non-linear functional analysis and applications, Edited by L.B. Rall. New York: Academic, 1971.

[19] J.M. Orega and W.C. Rheinboldt, Iterative solution of nonlinear equations in several variables. New York: Academic, 1970.

[20] M.M. Vainberg, Variational methods for the study of nonlinear operators. San Francisco: Holden Day, 1964.

[21] L. Collatz, Functional analysis and numerical mathematics. New York: Academic, 1966.

[22] T.L. Saaty and J. Bram, Nonlinear mathematics. New York: McGraw-Hill, 1964.

[23] L. Liusternik and V. Sobolev, Elements of functional analysis. New York: Frederick Ungar, 1961.

[24] A.N. Kolmogorov and S.V. Fomin, Functional analysis Vol. 1. Rochester, New York: Graylock, 1965.

[25] R.E. Edwards, Functional analysis theory and applications. New York: Holt Rinehart and Winston, 1965.

[26] L.V. Katorovich and G.P. Akilov, Functional analysis in normed spaces. New York: Macmillan, 1964.

[27] G.S. Christensen, "Aspects of nonlinear system stability". Doctoral Thesis, University of British Columbia, Vancouver, B.C., Canada, 1966.

[28] R. Courant and F. John, Introduction to calculus and analysis Vol. 1. New York: Interscience Publishers, .

[29] D.R. Lick, The Advanced calculus of one variable. New York: Appleton-Century-Crofts, 1971.

[30] K. Ogata, State space analysis of control systems. Englewood Cliffs, N.J.: Prentice Hall, 1967.

[31] D.G. Luenberger, Optimization by vector space methods. New York: Wiley, 1969.

[32] B.C. Kuo, Automatic control systems. Englewood Cliffs, N.J.: Prentice Hall, 1967.

[33] G.W. Trott and G.S. Christensen, "On the uniqueness of the Volterra Series". IEEE Transactions on Automatic Control, Vol. AC-14, No. 6, pp. 759-760, December, 1964.

[34] G.S. Christensen and G.W. Trott, "On the inclusion of initial conditions in Volterra Series". International Journal of Control, Vol. 12, No. 5, pp. 835-838, 1970.

[35] R.B. Parente, "Functional Analysis of systems characterized by nonlinear differential equations", Ph.D. dissertation, M.I.T. Cambridge, Mass., 1965.

[36] E. Hille and R.S. Phillips, Functional analysis and semigroups, Am. Math. Soc., 1957.

[37] D.A. George, "Continuous nonlinear systems", M.I.T. Electronics Research Lab., Cambridge, Mass., Rept. 355, 1959.

[38] R.S. Rao and G.S. Christensen, "On the convergence of a discrete Volterra Series". IEEE Transactions on Automatic Control, Vol. AC-15, No. 1, pp. 140-142, February, 1970.